INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

University of Alberta

Development and Application of Open Tubular and Open Tubular Capillary IMAC Devices for Phosphopeptide Enrichment and Characterization

by

Jacek Stupak

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science

Department of Chemistry

Edmonton, Alberta
Spring 2005

University of Alberta

Library Release Form

Name of Author: Jacek Stupak

Title of Thesis: Development and Application of Open Tubular and Open Tubular Capillary IMAC Devices for Phosphopeptide Enrichment and Characterization

Degree: Master of Science

Year this Degree Granted: 2005

Jan. 20, 2005

To all those who thirst for knowledge

By many ways, in many ways, I reached my truth:
It was not on one ladder that I climbed to the height
where my eye roams over my distance. And
it was only reluctantly that I ever inquired about
the way... I preferred to question and try out the
ways themselves. A trying and questioning was
my every move; and verily, one must also
learn to answer such questioning.
"This is my way; where is yours?"—thus I
answered those who asked me 'the way'.
For 'the way'—that does not exist.

—Friedrich Nietzsche

Abstract

Immobilized metal-ion affinity chromatography (IMAC) is a useful method for enrichment of phosphopeptides from enzymatic digests of proteins. Detection and identification of phosphopeptides is thereby greatly facilitated, since most non-phosphorylated peptides are eliminated from the resulting mass spectra. In this study, IMAC ligands were attached directly to the silanol groups of open glass tubing and fused silica capillaries. The resulting IMAC devices make for more effective and sensitive affinity capture than commercially available sepharose-based IMAC. The sensitivity advantage is especially pronounced for capillary-based IMAC, owing to the small elution volumes in the miniaturized IMAC device. Application of open tubular IMAC to the human Na^+/H^+ exchanger protein His182 allowed for mapping of 5 phosphorylation sites by MALDI-MS/MS. The miniaturized open tubular capillary IMAC has been shown to be able to detect phosphopeptides from gel-separated proteins at the sub-pmol level. Furthermore, it is also shown that following methylation of carboxylate residues in peptides, application of the negative-ion mode is an interesting way to selectively ionize phosphopeptides.

Acknowledgements

I would like to thank my supervisor Dr. Liang Li for his encouragement, guidance, financial support, and patience during the completion of this thesis. Also for giving me a chance to work in such a well-equipped, world-class laboratory.

I'd like to thank my committee members for reviewing and commenting on my thesis.

All members of the Li group are also gratefully acknowledged, in no particular order: Dr. N. Zhang , Mr. C. Ji , Dr. A. Doucette, Dr. D. Craft, Mr. Chris McDonald, Dr. H-Y. Zhong, , Ms. L. Tao, X. Yu, Dr. N. Li, Dr. Y. Zhang, Mr. Z. Wen, Mr. Bryce Young, Mr. M. Gebre, and Ms. N. Guo. Special thanks also to Dr. H. Liu, who worked on this project with me, for all his help and insightful comments; Also Dr. B. Keller for his expertise in capillary micro-spotting using a microscope in the beginning stages of this research.

Finally, thanks also go to my family for their support during my stay in Edmonton.

Table of Contents

List of Figures

List of Abbreviations

CID	Collission-induced dissociation
DHB	2,5-Dihydroxybenzoic acid
ESI	Electrospray ionization
HCCA	α-Cyano-4-hydroxycinnamic acid
IDA	Iminodiacetic acid (a.k.a. diglycine)
IMAC	Immobilized metal-ion affinity chromatography
MALDI	Matrix assisted laser desorption-ionization
MCP	Multichannel plate detector
MS	Mass Spectrometry
m/z	Mass-to-charge ratio
OT-IMAC	Open-tubular IMAC
OTC-IMAC	Open-tubular capillary IMAC
pM/pmol	Picomolar/picomole
Q-TOF	Quadrupole-time-of-flight
R_f	Radio frequency (AC)
TOF	Time-of-flight

Chapter 1

Introduction:

Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry and Immobilized Metal-Ion Affinity Chromatography in Phosphopeptide Characterization.

Protein identification and characterization by mass spectrometry (MS) has become one of the benchmark techniques in proteome analysis. Thanks to the development of soft ionization methods like matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI), which preclude thermal decomposition of large analytes like proteins, a constantly evolving, dynamic and multifaceted field has emerged. Methods for the analysis of post-translationally modified proteins by MS have received a great deal of interest recently, especially as new technical advances are constantly being made, gradually allowing these methods to become more reliable and robust. In the sections that follow, an overview of the MALDI ionization process and time-of-flight (TOF) as well as quadrupole-time-of-flight (QTOF) mass spectrometry will be given. We will also outline immobilized metal-ion affinity chromatography (IMAC) for phosphopeptide capture and enrichment, and give a brief introduction to protein separations by sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE).

1.1. MALDI-TOF Mass Spectrometry

1.1.1. Matrix-assisted laser desorption/ionization process

Matrix-assisted laser desorption/ionization (MALDI)[1-3] is a method for generating vapour-phase ions of large non-volatile polymers directly from a solid phase. For this purpose, the analyte is mixed with a matrix, typically a small organic compound, in an analyte to matrix ratio of greater than about 1:500. Upon crystallization, the analyte then becomes embedded in the crystal lattice of matrix molecules. The sample and matrix are introduced into a mass spectrometer's ion source region under vacuum, and upon application of a pulsed laser beam, both matrix and analyte molecules are desorbed and carried as a 'plume' into the gas phase. The matrix molecule is typically a small UV absorbing aromatic organic acid, with α-cyano-4-hydroxycinnamic acid (HCCA) and 2,5-dihydroxybenzoic acid (DHB) being the most common. The main function of the matrix is to absorb energy from the laser, whereby the molecules become electronically excited, and then transfer some of this energy onto analyte molecules, facilitating their desorption. Since the matrix is present in large excess compared to the analyte, isolation of polymer molecules takes place, where formation of analyte aggregates is prevented, analyte-target interactions are minimized, and sample decomposition is avoided[4,5]. Finally, the matrix also acts as a Bronsted acid or base, donating or accepting a proton from the analyte in order to ionize it.

Since sample and matrix preparation is crucial to a successful MALDI-MS analysis, several sample preparation methods have been developed, including

the dried-droplet[2], vacuum-drying[6], crushed-crystal[7], fast evaporation[8], sandwich[9] and two-layer[10]. In our laboratory, the most often used method for MALDI-MS sample preparation is the two-layer method with HCCA matrix. In this method, a first thin layer of HCCA matrix is deposited on a steel target; a solution of analyte in a saturated matrix solution is then applied on top of the first layer. The first layer acts as a seed for crystallization, giving uniform, homogeneous crystals which can then be washed with water to remove ionic contaminants or salts. The two-layer method provides high-resolution, sensitivity, and has excellent shot to shot reproducibility, owing to the absence of 'sweet spots' (regions of high analyte concentration) that we would normally observe with any single-layer method. This method has also been shown to be ideal for samples containing up to 1% SDS[11].

Pulsed lasers in UV-MALDI are typically nitrogen or Nd:YAG, while CO_2 or Er:YAG lasers have been used in IR-MALDI. Figure 1.1 shows a simplified schematic of the MALDI process. Following the impact of the laser, a 'plume' of analyte and matrix molecules is transferred into the gas phase under the vacuum conditions typically used in MALDI. This 'plume' contains an assortment of neutral, protonated, and free-radical species. The exact mechanism in MALDI is still unclear, and several models for the ionization process have been proposed. Regardless of whether proton transfer takes place in the gas or solid phase, application of a short positive voltage to the resulting gas-phase ions so generated then allows us to accelerate and detect only positively charged species

(protonated or cationized analytes); application of a short negative voltage allows us to accelerate and detect only negatively charged ions (deprotonated analytes; negative-ion MALDI-MS).

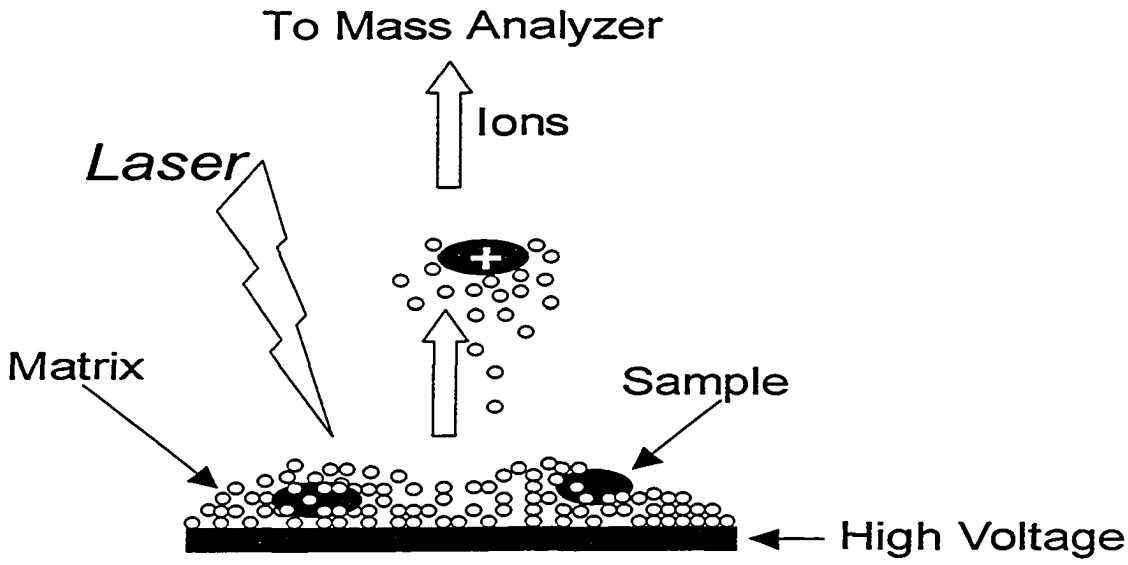

Figure 1.1. The process of matrix-assisted laser desorption ionization (MALDI).

1.1.2 Time-of-Flight Mass Analyser

A number of mass spectrometers have been used to analyze the ions produced by MALDI, including Fourier-transform ion cyclotron resonance[12], magnetic sectors[13], and quadrupole ion traps[14]; however, the time-of-flight mass analyzer remains as the most commonly used instrumental setup, owing to its simplicity, sensitivity, and virtually unlimited mass range. At the heart of the time-of-flight analyzer (Fig.1.2) are the source-extraction region where ionization takes place, the field-free drift region where ions fly towards the detector, and the high-vacuum state of the entire system.

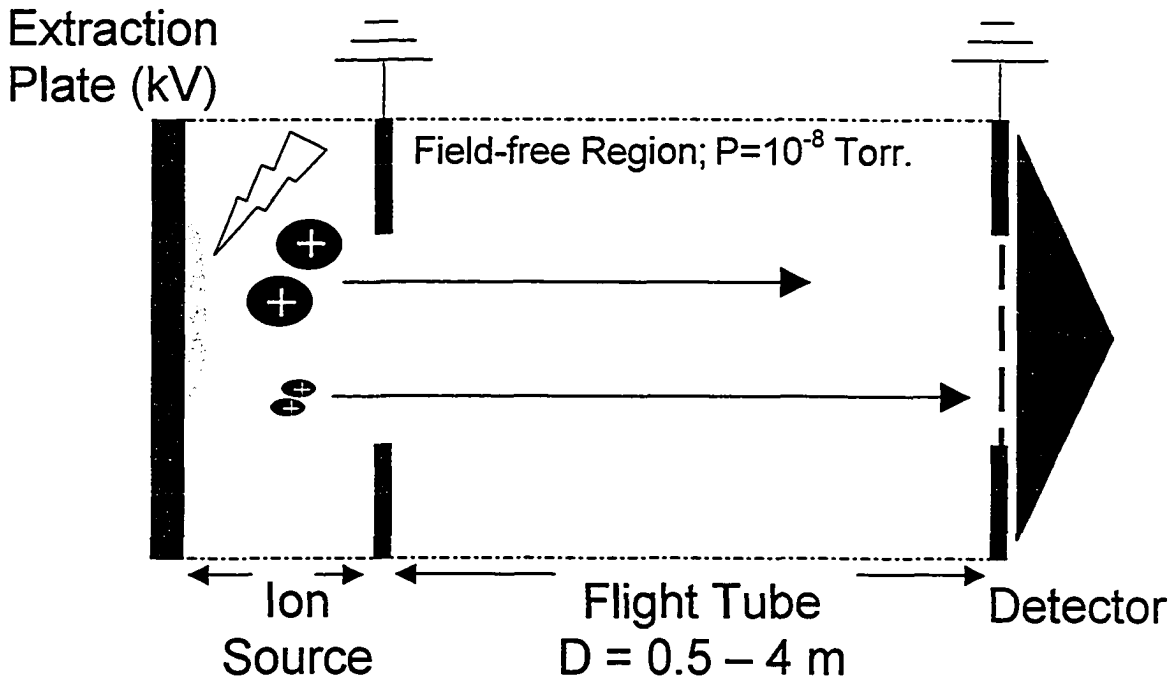

Figure 1.2. The Time-of-Flight (TOF) Mass Analyzer.

As analyte ions enter the gas phase in the source-extraction region through the MALDI process, a voltage is used to accelerate these ions through a long field-free drift region towards the detector. The ions travel with different velocities depending on their mass-to-charge ratios; hence measurement of the 'time-of-flight' of an ion through the drift tube can then be used to calculate the mass for that ion.

When a packet of ions with mass m is accelerated by application of a voltage V, the ions travel with a velocity v, expressed by

$$Ve = \tfrac{1}{2} mv^2 \text{ (eq. 1.1.1)}$$

where e is the charge of an electron. This equation is obtained from the law of conservation of energy, as the total energy from the accelerating voltage (left side of the equation) must equal the kinetic energy imparted on the ions (right side of

the equation). The above equation can be rearranged to $v = (2Ve/m)^{1/2}$, showing that velocity is inversely proportional to mass, and that the ions travel with a velocity proportional to $(m/z)^{-1/2}$. Put simply, from the typically singly charged ions obtained with MALDI, ions with lower masses will travel faster through the drift tube and hence have a shorter time-of-flight than larger ions. If we then substitute the equation for velocity into eq. 1.1.1, we can then obtain an equation describing the relationship of an ion's mass to its flight time.

$$\text{Velocity} = \text{Distance} / \text{Time}; \quad v = D/t \quad \text{(eq.1.1.2)}$$

Substituting 1.1.2 into 1.1.1 gives,

$$t = (mD^2/2Ve)^{1/2} \quad (1.1.3)$$

Equation 1.1.3 shows that if we take the accelerating voltage and flight distance to be a constant A, then the time-of-flight of an ion can be described by $t = A\,(m/z)^{1/2}$; in other words it is proportional the square root of the mass-to-charge ratio of that ion. To account for instrumental imprecision and experimental variance, a TOF mass spectrometer is further calibrated with known standards to allow accurate mass measurement.

1.2. Mass Resolution in a TOF Mass Analyzer

Mass resolution (R) in a TOF mass analyzer is determined by the peak width at a given mass in a mass spectrum:

$$R = m/\Delta m \text{ (eq. 1.2.1)}$$

In terms of the flight of an ion through the field-free drift tube, the above equation can be expressed as

$$R = t/2\Delta t \text{ (eq. 1.2.2)}$$

Where t is the peak width expressed as the full width at half maximum (FWHM) of the signal, the peak width at half height. Ideally, discrete packets of ions with different m/z ratios would arrive at the detector at different times, and these times would be exactly the same for all ions of the same m/z ratio. However, on account of the spatial distribution of ions as well as their initial kinetic energy distribution, ions of the same m/z ratio typically arrive at the detector spread out in a finite time window, which leads to each peak for a particular ion having a given width, and limits the resolution for the analysis[15-17]. Hence, minimizing the spread of the space and kinetic energy distributions gives sharper peaks with better resolution.

Since the initial spatial spread of ions in MALDI is typically very small on account of the thin layer of solid matrix/analyte spotted on a flat metal surface, minimizing the initial kinetic energy distribution has given rise to the modern state-of-the-art MALDI-MS instruments in use today. These high-resolution instruments use time-lag focusing of ions, and ion mirrors to improve resolution, and are discussed in the following sections.

1.2.1. Time-lag focusing, or Delayed Extraction[18-20].

Fig. 1.3 illustrates the concept of time-lag focusing. Following formation of ions by the MALDI process in the source region of the mass spectrometer, ions are allowed to expand into the source region of the mass spectrometer. Since both the repeller and first extraction grid are held at the same potential (Fig. 1.3A), for ions of the same m/z ratio, the ions with higher initial kinetic energies will move farther away from the repeller. After a short delay, an extraction pulse is superimposed on the repeller plate, which extracts the ions into the flight tube (Fig. 1.3B). Because less energetic ions find themselves closer to the repeller than more energetic ions at the start of this pulse, they end up experiencing more accelerating energy; this in turn allows them to catch up to the more energetic ions on their way to the detector (Fig.1.3C). This method greatly reduces peak broadening, and leads to packet of ions with the same m/z ratio to arrive at the detector with a greatly reduced time distribution window.

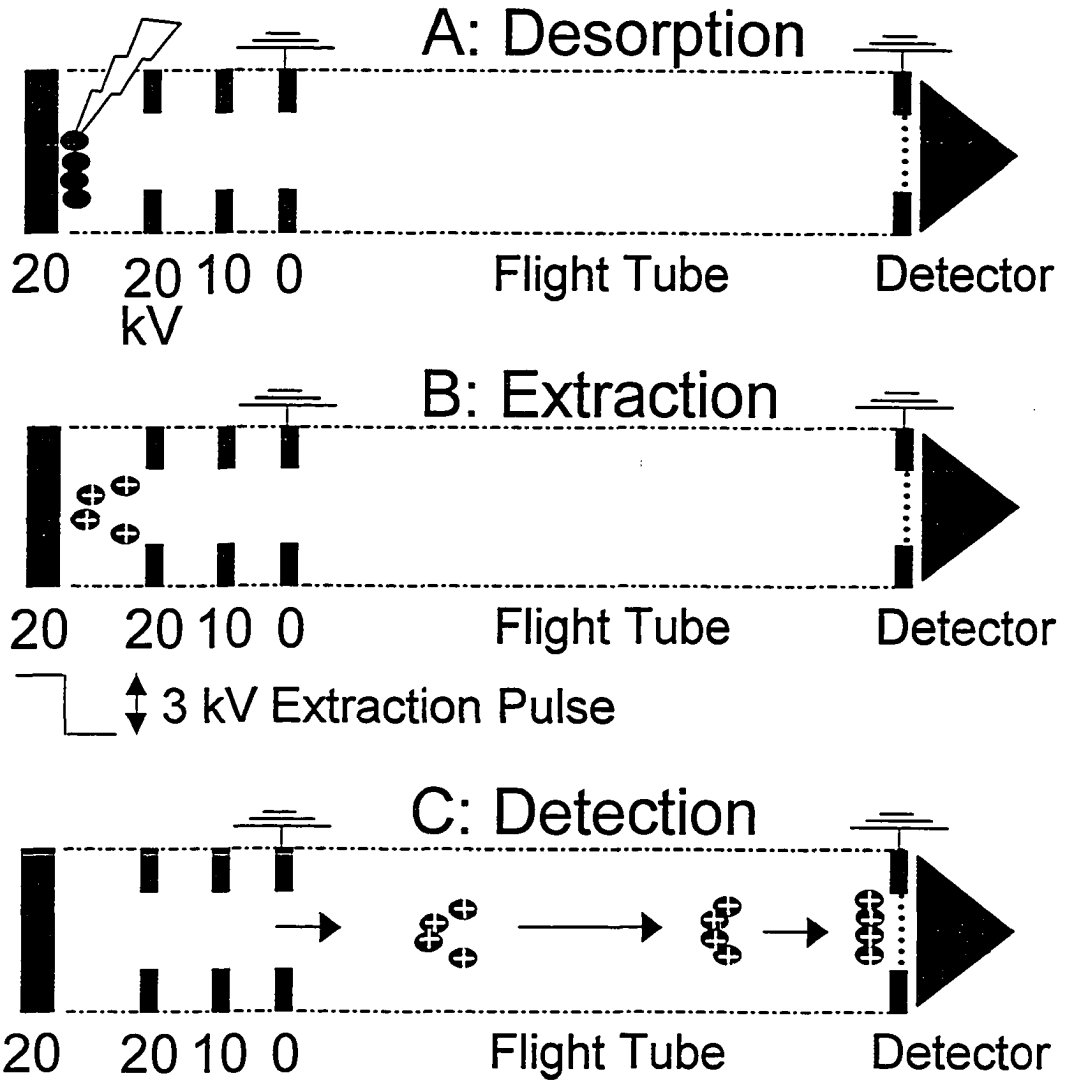

Figure 1.3 The Concept of Time-Lag Focusing.

1.2.2. Reflectron or Ion Mirrors to Improve Resolution

A linear TOF instrument such as that represented by the simple setup in Fig.1.1.1 is known a low-resolution mass spectrometer. A setup with a reflectron[21-22], shown in Fig. 1.4, incorporates a number of ion mirrors to reflect ions in the direction of the detector. The ion mirror consists of a series of electric plates, with gradually increasing applied potentials. For ions of the same m/z ratio, the more energetic ions will move faster and hence penetrate deeper into the

electric field produced by the mirrors; they will hence spend more time in the reflectron than less energetic ions of the same m/z ratio. As these ions penetrate into the mirrors, they will then experience a greater applied voltage, and will be focused in time with the other ions of the same m/z ratio at the detector. This process, like time-lag focusing, minimizes the kinetic energy distributions of ions, leads to sharper peaks, and thus improves resolution.

Modern TOF-MS instruments typically incorporate both time-lag focusing and a reflectron, and are considered to be capable of high-resolution. With improvements in resolution, these instruments produce sharp and narrow peaks, which also leads to greater sensitivity because of the concomitant increase in peak height with decreased peak width. In addition, the TOF itself is a very fast and sensitive mass analyzer owing to its high ion transmission (i.e., all ions of different m/z ratios can be detected in one ionization pulse), versus a scanning instrument like a quadrupole, which must transmit one ion of a given m/z ratio at a time.

Although theoretically limitless, the mass range in a TOF instrument is typically limited by the effectiveness of the detector and acceleration voltages and the maximum m/z range is about 1 million[23].

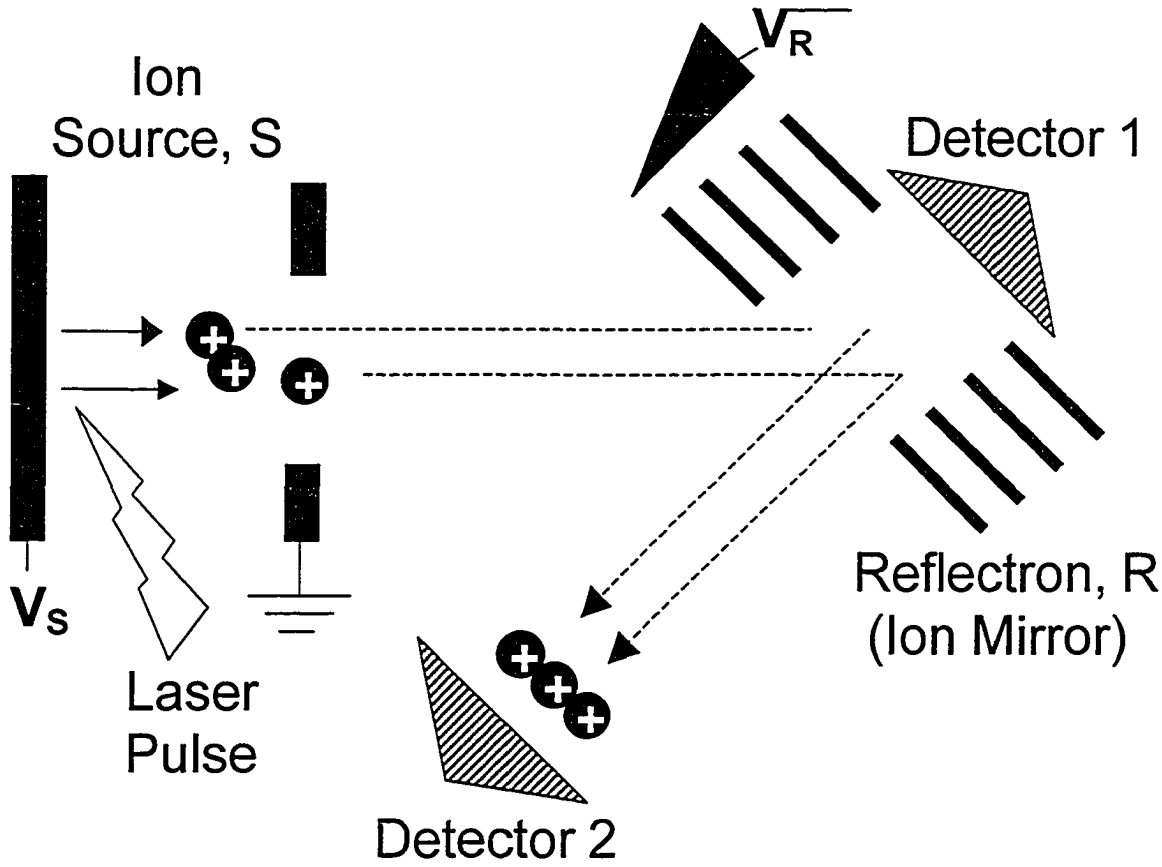

Figure 1.4. Schematic of a Reflectron, or Ion Mirror Used to Improve Resolution in MALDI-TOF-MS.

1.3. Detector

A commonly used detector in TOF instruments is the multi-channel plate detector, or MCP[23-25], shown in Fig. 1.5. It is a thin disc with millions of small channels with lead-doped glass. Each channel acts like an electron multiplier: for each ion striking the surface of the channel, a cascade of electrons is released and multiplied as the electrons travel towards the anode. The ion current is then digitized into an MS signal.

Since high-mass ions travel with lower velocities, their impact on the detector produces a less powerful electron cascade than would be the case for a smaller, faster ion. To increase detection sensitivity of large polymers, high accelerating voltages must be used, or alternatively, post-acceleration in the flight-tube or detector region has also been employed.

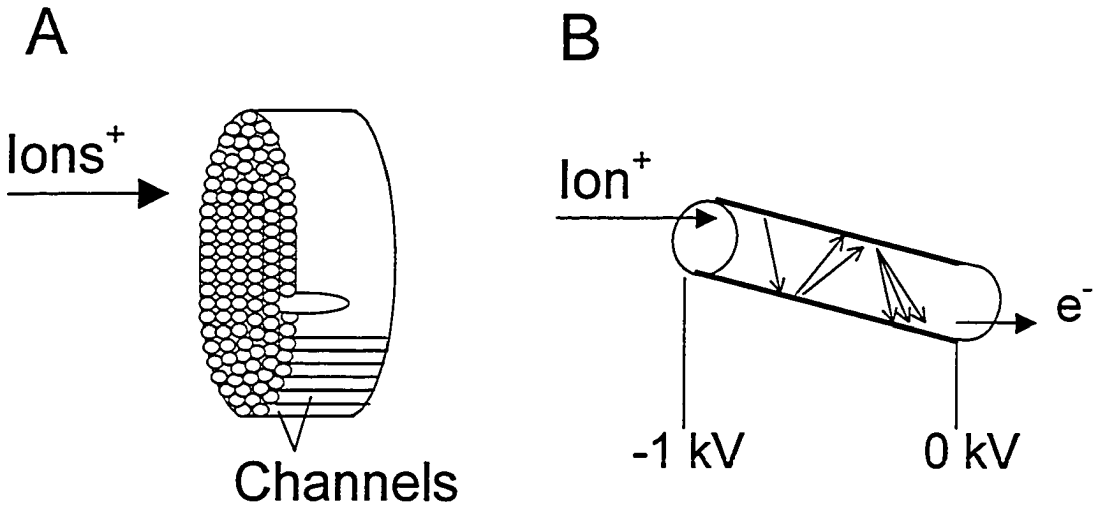

Figure 1.5. Schematic of a Multi-Channel Plate Detector used in TOF instruments. A: The MCP contains millions of lead-doped channels of glass. B: Impact of a flying ion on the surface leads to a cascade of electrons towards the anode.

1.4. Tandem Mass Spectrometry: The Quadrupole-TOF Mass Analyzer

Once the mass of a compound has been assigned by mass spectrometry, one would then like to fragment the given compound in order to obtain structural and/or sequence information. This is typically done by connecting two mass spectrometers in tandem. Molecules exiting the first mass analyzer enter a cell where they collide with gas molecules, gives rise to product ions, fragments which are then separated and detected by the second mass analyzer. Various

combinations of mass analyzers have been utilized in tandem-MS instruments, including triple-quadrupole, sector-sector, ion-trap-TOF, TOF-TOF, and quadrupole-TOF. Each instrument gives rise to characteristic collision-induced dissociation (CID) fragmentation patterns, and each has its own distinct advantages. In our laboratory we currently use a quadrupole-TOF instrument fitted with a MALDI source, and we will explain its functioning in the following section.

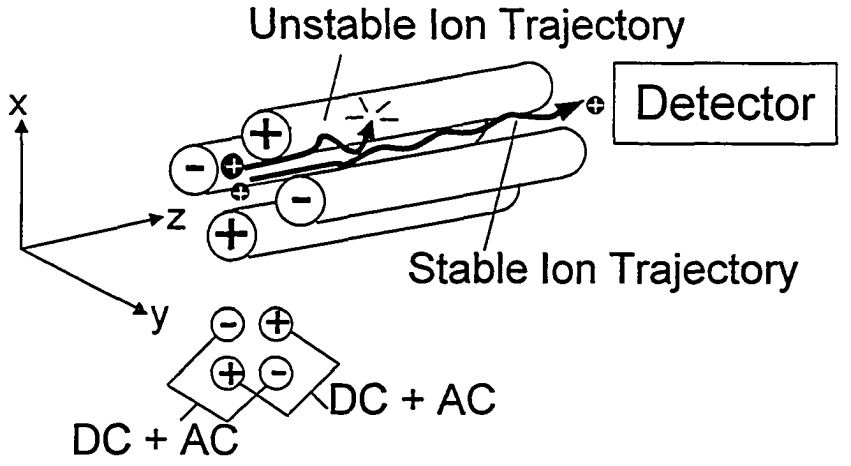

Figure 1.6. The Quadrupole Mass Spectrometer.[25]

The quadrupole mass analyzer (Fig. 1.6) consists of four cylindrical rods, electrodes to which both an AC and a DC voltage are applied. One pair of rods is connected to a positive DC source, the other to a negative, and the applied AC radio frequency signals are 180 degrees out of phase. Ions are first accelerated into the quadrupole by a small potential of 5 – 10 V. At a particular value of the AC and DC voltages, only ions with a given m/z ratio will pass through the quadrupole to the detector, while all other ions will oscillate into the rods and be annihilated. Thus, when we simultaneously increase the AC and DC voltages

while keeping their ratio constant, we can then scan the mass range and obtain a mass spectrum[27-28].

The potential at any point (x, y) in the hyperbolic field created by the four rods of a quadrupole is defined by the Mathieu equations:

$$\Phi = [\, U + V\cos(wt)\,]\, x^2 - y^2 / 2r_o{}^2 \quad \text{(eq. 1.3.1)}$$

Where U represents the magnitude of the applied DC potential, the Vcos(wt) term stands for the magnitude and frequency of the applied AC voltage, and r_o is the distance from the center of the field between the rods to the nearest surface. By a set of complex mathematical operations, equation 1.3.1 can be rearranged and simplified into

$$d^2U/d\xi^2 + [a_u + 2q_u\mathrm{Cos}2\xi]U = 0 \quad \text{(eq.1.3.2)}$$

Where $a = 4eU/w^2r_o{}^2m$ and $q = 2eV/w^2r_o{}^2m$. Defining the precise solutions to equation 1.3.2 is far beyond the scope of this work . However, qualitatively, the solutions can be classified as bounded or unbounded solutions. Bounded solutions correspond to cases where the trajectory of the ions in the x and y directions would be finite as it moves along the z axis towards the detector. Unbounded solutions correspond to cases where the trajectory of the ion is unstable, causing it to crash on the way to the detector. Plotting different values of a and q where solutions to equation 1.3.2 are stable or unstable gives the a-q stability diagram[27-28]. This diagram, shown in Fig. 1.7, allows easy access to understanding the resolution capabilities of a quadrupole mass spectrometer.

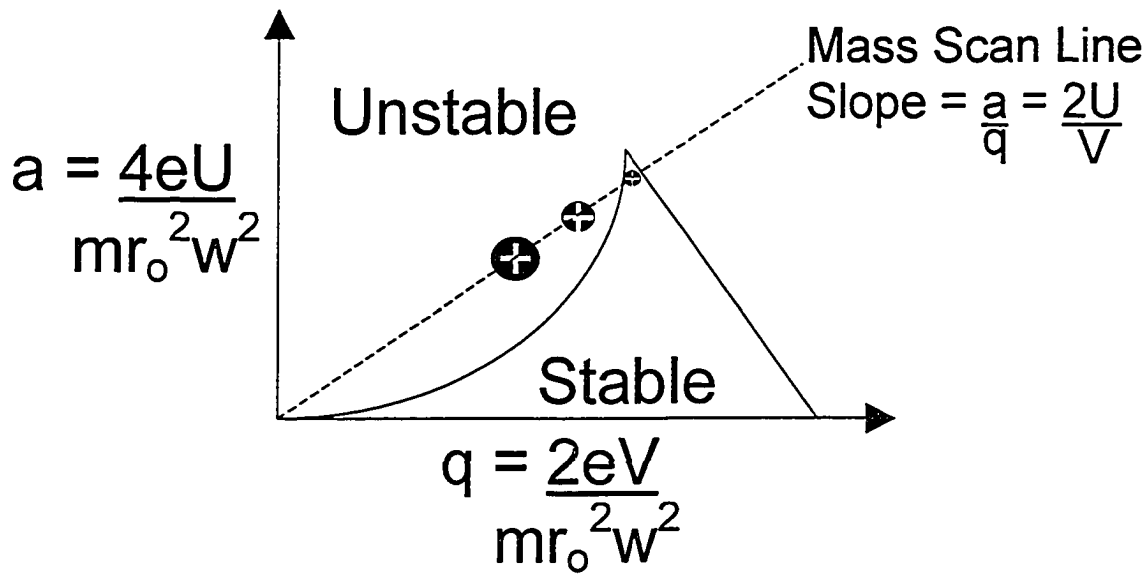

Figure 1.7. The a-q stability diagram.

Quadrupole instruments are operated by simultaneously increasing the AC and DC voltages while keeping their ratio constant; in terms of the stability diagram shown above, this means that the ratio of a to q remains the same while their absolute magnitudes vary. Plotting these values gives the mass scan line with slope = a/q = 2U/V, and zero intercept. As we increase the magnitude of both U and V while keeping the 2U:V ratio, the mass which appears at the tip of the stable portion of the a-q diagram will increase correspondingly. The slope of the mass scan line shown in Fig. 1.7 represents a case optimized for mass resolution, as the scan line crosses the sharp and narrow region of the a-q stability diagram. When the slope of the mass scan line is lowered, that is when the magnitude of the DC potential is lowered relative to the AC potential, the quadrupole then transmits a wider window of m/z ratios in a single scan. When the DC potential is 0, the mass scan line in Fig. 1.7 then runs horizontally along the x axis, and the

quadrupole becomes a RF-only broad bandpass mass filter, which transmits ions having m/z ratios higher than the cut-off value corresponding to q = 0.908. The mass range of quadrupole instruments is limited by the magnitude of the applied voltages, and in modern instruments is typically 3000-4000 Da.

1.4.1 Quadrupole-TOF[30] and MS/MS

Fig. 1.8 shows a simplified schematic of a quadrupole-TOF instrument. This instrument is able to acquire both MS and MS/MS fragment ion spectra. When operated in MS-only mode, the 2 quadrupoles are operated as RF-only broad bandpass mass filters, while the sensitivity and resolution of the TOF are taken advantage of in measuring the mass of the ions being analyzed. In acquiring MS/MS fragment ion spectra for a compound of interest, Q1 operates as a narrow-bandpass mass filter, and transmits the ion of a given m/z ratio at one time. The ion is reaccelerated as it enters the Q_2 quadrupole, where it collides with gas molecules like Ar or N_2, and undergoes collision-induced dissociation, giving rise to fragment ions of the parent compound. These fragment ions are focused by ion optics and reaccelerated into the TOF mass analyzer equipped with time-lag focusing and an ion mirror. Interpretation of the MS/MS spectrum then allows assignment of the parent ion identify and chemical structure.

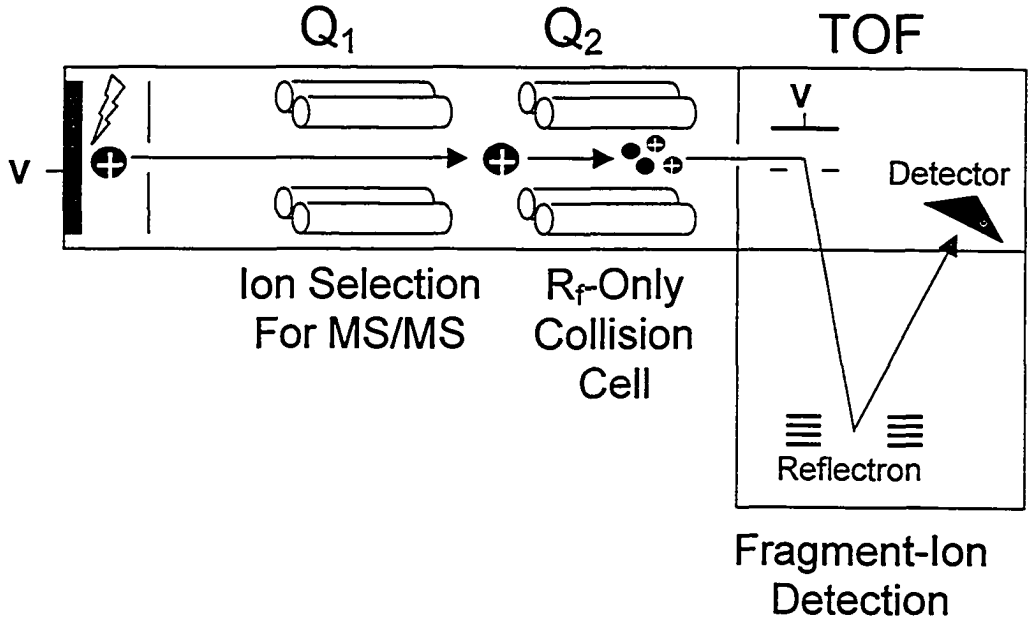

Figure 1.8. Quadrupole-TOF Mass Spectrometer.[29]

1.5. MS/MS Fragmentation and Interpretation of Spectra[31-34].

MS/MS of peptides with a quadrupole-TOF instrument gives rise to fragments by low energy collision-induced dissociation. The distinction between low and high energy CID is based on the potential used to accelerate ions into the mass spectrometer, it being in the eV range for low-energy CID, and in the KeV range for high-energy CID. High energy CID is performed on older four-sector instruments as well as in newer TOF-TOF instruments. Because a low-energy CID cell is an RF-only quadrupole, this mode has a high ion transmission efficiency and a high fragmentation efficiency owing to the long collision path in the quadrupole, which gives rise to multiple collisions with gas molecules. In general, large peptides are more difficult to fragment than smaller ones, having more degrees of vibrational freedom over which to distribute the kinetic energy

transferred as a result of collision with gas molecules in the collision cell. With our quadrupole-TOF instrument, the upper mass range limit is 3000 Da.

The nomenclature of MS/MS fragmentation was first put forth by Roepstorff and Fohlman[31] and has been reviewed. Fig.1.9 shows the nomenclature for typically observed low-energy CID fragments from peptides. A number of parameters influence the patterns and types of fragmentation, including collision gas pressure, type of collision gas, the effective charge of the ion, and the amino-acid sequence of the peptide.

Only fragments with a positive charge will be detected by the second mass spectrometer following CID, as neutral molecules and radicals will be lost. When the positive charge of the parent peptide is retained on the N-terminal fragment following CID, the resulting ions are classified as being of the a, b, or c type. In this case, the fragment originating from the C-terminus forms a neutral molecule and is lost. On the other hand, when the charge is retained on the C-terminal fragment of the parent peptide, the ions are then termed as being of the x, y, z type and the neutral fragment from the N-terminal side is lost.

In addition to the N or C terminal ions shown in Fig. 1.9, double cleavage of the peptide backbone can also give rise to internal fragment ions. These are shown in Fig. 1.10. They can result from b-type and y-type cleavage to produce amino-acylium ions; a-type and y-type cleavages give rise to immonium ions. Losses of H_2O (-18 Da) and NH_3 (-17 Da) are from peptides and fragments are also common; furthermore, losses of 98 or 80 Da usually indicate the presence of

phosphorylation, as they correspond to losses of H_3PO_4 and HPO_3, respectively. The presence of phosphorylated immonium ions of tyrosine at m/z 216 in combination with a neutral loss of 80 or 98 Da can also indicate the presence of phosphotyrosine. Interestingly, although rare, tyrosine sulfation also results in neutral losses of 80 Da upon CID analysis[34].

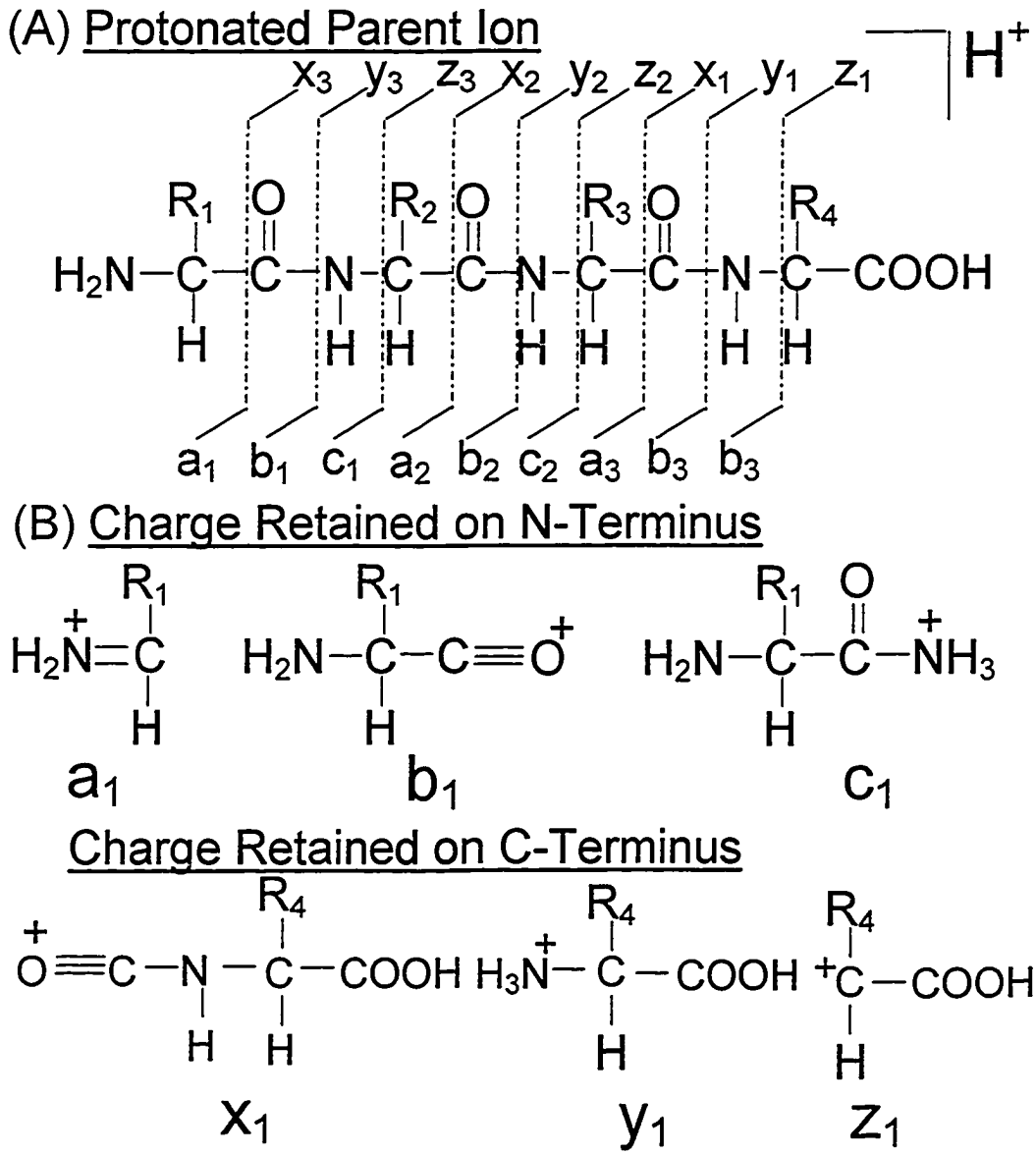

Figure 1.9 Nomenclature of low-energy CID, MS/MS fragmentation of peptides. A: Schematic of how a, b, c, x, y, and z ions form upon collision-induced dissociation in a tandem mass spectrometer. B: Peptides resulting from the cleavages shown in A.

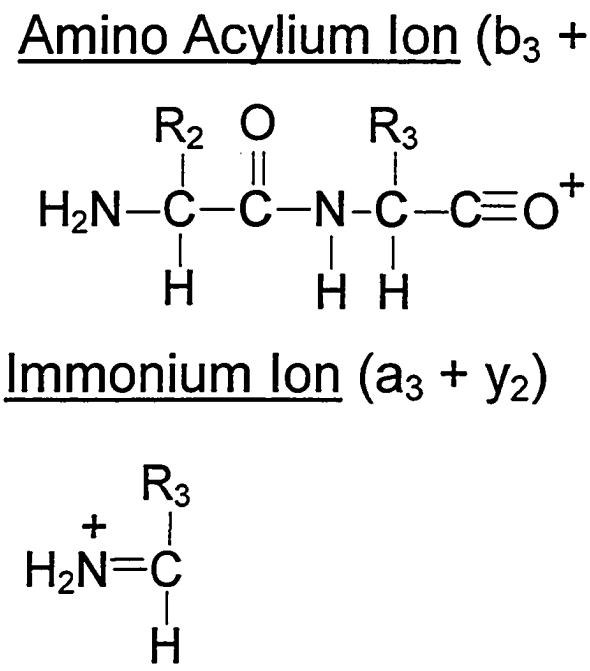

Figure 1.10. Internal fragment ions.

Peptides generated by MALDI quadrupole-TOF MS/MS often show preferential cleavage on the C terminal side of aspartic and glutamic acid residues, as well as on the N-terminal side of prolines. In contrast, C-terminal cleavage on prolines is often absent. Because the retention of charge of a fragmented peptide is in some measure determined by the ability of particular amino acids to transfer and retain protons, observation of b or y-type ions on fragments containing the basic residues lysine, arginine, histidine, tryptophan and proline has also been reported[35].

In general, interpretation of MS/MS spectra is seldom done in a *de novo* manner. Having obtained a good quality MS/MS spectrum, one typically feeds the results into an internet program such as Protein Prospector Mass-Tag, or Mascot.[36-38] Taken together with peptide mass mapping results from MS-only

identification of tryptic peptides of the protein, one can then gain confidence in the protein identity assignment. Assignment and interpretation of MS/MS spectra for a post-translationally modified protein such as a phosphoprotein is more complex, as the location of the phosphate group is often the goal of the analysis. Assignment of phosphorylation sites is often an arduous task involving multiple analysis and more than one endoproteinase, since unambiguous phosphorylation site assignment requires us to observe given fragments, which unfortunately are sometimes not observed. This problem is often further complicated by inefficient fragmentation of large peptides, and ion suppression of low abundance phosphopeptides by more numerous non-phosphorylated species.

1.6. Immobilized Meta-Ion Affinity Chromatography (IMAC)

A convenient way to facilitate the detection of low-abundance and/or poorly ionizing phosphopeptides by mass spectrometry is to selectively capture and preconcentrate them whilst removing the background of non-phosphorylated peptides. The affinity of Fe(III) for phosphate was first recognized by Andersson and Porath in 1986[39], and since then a number of commercial products for phosphopeptide affinity capture appeared in the 1990's. As shown in Fig.1.11, IMAC consists of an immobilized ligand such as diglycine, which chelates and in turn immobilizes a metal-ion such as Fe(III), for example. The chelated metal is then free to form a coordination complex with phosphate. The strengh of the resulting bond is highest between pH 3 and 5 owing to the second ionization constant of phosphate, and decreases at alkaline values due to for

binding with hydroxyl ions. The tridentate diglycine ligand shown in Fig. 1.11 can be one of a number of structures, including tridentate phosphoserine, bidentate 8-Hydroxyquinoline or hydroxamic acid, and the tetradentate nitrilotriacetic acid or tris(carboxymethyl) ethylenediamine[40]. A spacer is usually included between the ligand and immobile support to which it is attached. Numerous supports have been utilized to immobilize the said ligands, the most popular of these being carbohydrate base polymers such as sepharose and agarose derivatives, and synthetic polymers such as polystyrenedivinylbenzene or poly(hydroxy)-methacrylate beads[41].

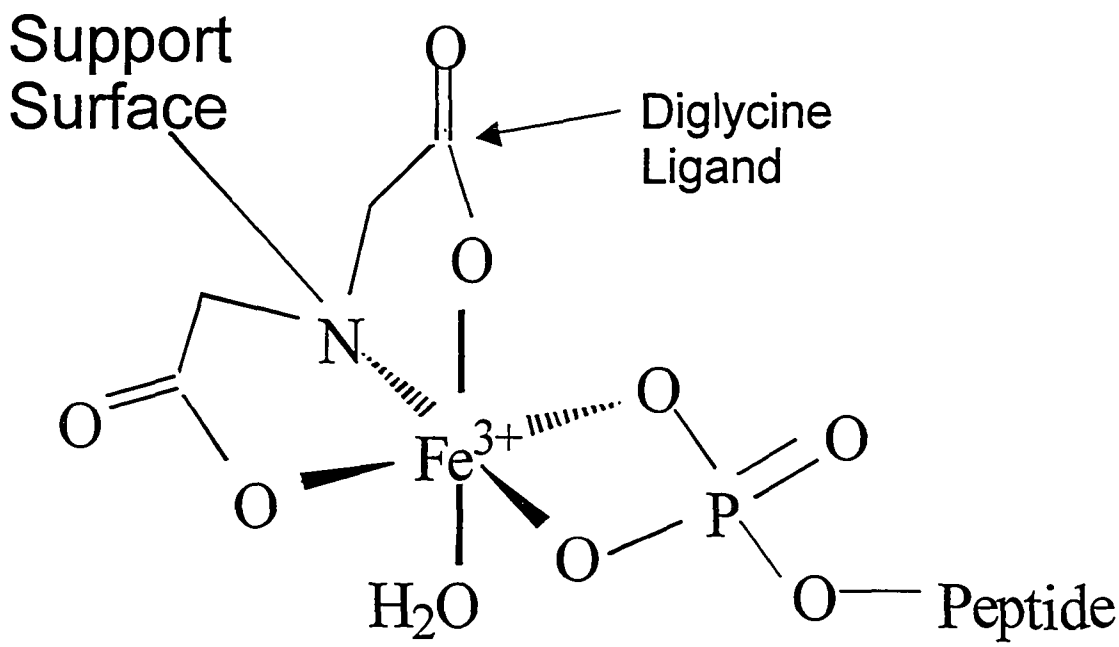

Figure 1.11. Coordination complex formed by immobilized diglycine-Fe(III) with a hypothetical phosphopeptide.

Cu(II) and Ni(II) IMAC has been used extensively in biochemistry for purification of recombinant proteins bearing a polyhistidine tag, and also shows affinity towards tryptophan and cysteine[40, 42]. Fe(III), Ga(III), Zr(IV), Al(III)

IMAC on the other hand show increased affinity for phosphate, and to a lesser extent also carboxylate, with Fe(III) and Ga(III) IMAC usage being most common. In carrying out IMAC enrichment of phosphopeptides, non-specific binding through the carboxylate groups can be decreased by extensive washing following binding; however, the extent of non-specific binding appears to be largely sample dependent. Recently, the notion of non-specific binding occurring through the carboxylate group has been challenged[43]. In analyzing the entire plasma membrane phosphoproteome en masse in Arabidopsis, Nuhse *et. al.* observed no correlation between peptides' physicochemical properties and number of acidic residues, with non-specific binding. Elution of phosphopeptides from IMAC can be done by increasing the pH (ex./ NH_4OH); however using free phosphate to achieve complete elution of multiply phosphorylated peptides has also been emphasized[44].

Although the total number of IMAC products on the market is high, using IMAC for phosphopeptide enrichment and characterization is still considered by some as "unreliable for 'real life' proteins"[45,46]. Recently, a survey test-sample was sent out to labs around the world to gauge the effectiveness of existing methods for phosphorylation site mapping. The authors concluded that "the relative lack of success using IMAC enrichment suggests that optimized and well-characterized procedures for this approach still are lacking or not sufficiently disseminated among the scientific community."[47]

1.7. SDS-PAGE and Protein Characterization

Sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE)[48] is still one of the most widely used techniques for achieving protein separations. Its resolution, sample loading capacity, ability to quantitate stained protein spots, and wide dynamic range, combined with well established mass spectrometric methods for protein identification, have made it a staple of most proteomic efforts. In this technique, a sieving, porous gel is formed by the co-polymerization of acrylamide (CH_2=CH-NH_2) with the cross-linker bisacrylamide (CH_2=CH-CO-NH-CH_2-NH-CO-CH=CH_2) by way of a free-radical vinyl addition. Proteins boiled with the strong anionic detergent sodium dodecyl sulfate and treated with a disulfide-reducing agent like dithiothreitiol are denatured, losing their three dimensional structure and any disulfide-bond linkages. SDS then binds to proteins in a 1.4:1 g ratio, approximately 1 SDS molecule for every 2 amino acids. SDS-treated proteins acquire an overall negative charge, uniform shapes, and approximately equal mass-to-charge ratios. When electrophoresis is performed on SDS-treated proteins applied to a polyacrylamide-bis gel, proteins migrate towards the anode with the same effective force, owing to their nearly equal mass-to-charge ratios; however, the retarding force they experience from the porous gel is proportional to protein size, and a plot of distance traveled on the gel versus the log of MW shows an approximately linear relationship. Proteins are visualized by staining with a dye such as Coomasie-blue or metals such as silver or copper, with detection limits

being typically in the 20-100 ng range for Coomasie-blue and ~1-10 ng for silver. Resolution of highly complex protein mixtures by gel electrophoresis can be further improved by coupling SDS-PAGE in one dimension with isoelectric focusing of proteins in another[49]. In this way, proteins are focused into narrow zones based on charge in the first dimension, and are then separated by size in the second dimension. Maps comprised of up to 5000 proteins can be achieved using 2-dimensional SDS-PAGE[50].

Characterization of proteins separated by SDS-PAGE involves in-gel digestion of the protein into polypeptides using an endoproteinase. The enzyme trypsin, which cleaves proteins on the C-terminal sides of Arg and Lys residues, is most often used for this purpose. Peptides are then extracted from the gel, desalted, and analyzed by mass spectrometry. By matching several peptides to a given protein in a public protein sequence database using probability scoring then allows for tentative identification. MS/MS analysis and fragment-ion database searching on any number of the peptides then further increases the confidence of the protein assignment. Theoretical fragmentations for a peptide are easily generated[38] for manual spectrum inspection and sequencing, as in phosphorylation site assignment, for example.

1.8. References Cited

1. Karas, M., Bachmann, D., Bahr, Y., Hillenkamp, F. **1987**, *Int. J. Mass Spectrom. Ion Processes*, 78, 53-68.
2. Karas, M, Hillenkamp, F. **1988**, *Anal. Chem.*, 60, 2299-2301.

3. Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T. **1988**, *Rapid Commun. Mass Spectrom.*, 2, 151-153.

4. Hillencamp, F, Karas, M, Beavis, R.C., Chait B.T., **1991**, *Anal. Chem.* 63, 1193A-1202A

5. Karas, M., Gluckmann, M., Schafer, J. **2000**, *J. Mass Spectrom.*, 35, 1-12.

6. Weinberger, S.R., Boernsen, K.O., Finchy, J.W., Robertson, V., Musselman, B.D., **1993**, *Proceedings of the 41st ASMS Conference on Mass Spectrometry and Allied Topics*, San Francisco, CA, pp. 775a-b.

7. Xiang, F., Beavis, R.C., **1994**, *Rapid Commun. Mass Spectrom.*, 8, 199-204.

8. Vorm, O., Roepstorff, P., Mann, M., **1994**, *Anal. Chem.*, 66, 3281-3287.

9. Li, L., Golding, R.E., Whittal, R.M., **1996**, *J. Am. Chem. Soc.* 118, 11662-11663.

10. Dai, Y., Whittal, R.M., Li, L., **1996**, *Anal. Chem.*, 2494-2500.

11. Zhang, N., Doucette, A., Li, L., **2001**, *Anal. Chem*, 73, 2968-2975.

12. Hettich, R.L, Buchanan, M.V., **1991**, *J. Am. Soc. Mass Spectrom.*, 2, 22-28.

13. Strobel, F.H., Solouki, T., White, M.A., Russel, D.H., **1991**, *J. Am. Soc. Mass Spectrom.*, 2, 91-94.

14. Chambers, D.M., Goeringer, D.E., McLuckey, S.A., Glish, G.L., **1993**, *Anal. Chem.*, 65, 14-20.

15. Cotter, R.J., *Time-of-Flight Mass Spectrometry: Instrumentation and Applications in Biological Research*; American Chemical Society, Washington, D.C., 1997.

15. Beavis, R.C., Chait, B.T., **1991**, *Chem. Phys. Lett.*, 181, 479-484.

16. Cotter, R.J., **1992**, *Anal. Chem.*, 64, 1027A.

17. Wiley, W.C., McLaren, I.H., **1955**, *Rev. Sci. Instrum.*, 26(12), 1150-1157.

16. Whittal, R, Li, L., **1997**, *American Laboratory*, ???, ??, 30-36.

20. Spengler, B., Cotter, R.J., 1990, *Anal. Chem.*, 62, 793-796.

21. Mamyrin, B.A., Karataev, V.I., Shmikk, D.V., Zagulin, V.A., 1973, *Sov. Phys – JETP*, 37(1). 45-48.

22. Maryrin, B.A., **1994**, *Int. J. Mass Spectrom. Ion Processes*, 131, 1-19.

23. Yost, R.A., Enke, C.G., **1978**, *J. Am. Chem. Soc.*, 100, 2274-2275.

24. Evans, S., Methods in Enzymology: Mass Spectrometry, McCloskey, J.A., Ed., Vol. 193, Academic Press, San Diego, CA, 1990, 61-86.

25. Wurz, P., Gulber, L., **1996**, *Rev. Sci. Instrum.*, 67, 1790-1793.

26. Adapted from: Skoog, D.A., Holler, F.J., Nieman, T.A., *Principles of Instrumental Analysis*, 5th Edition, Harcourt Brace & Company, Toronto, 1998, 258.

27. Skoog, D.A., Holler, F.J., Nieman, T.A., *Principles of Instrumental Analysis*, 5th Edition, Harcourt Brace & Company, Toronto, 1998, 258-261.

28. Miller, P.E., Denton, M.B., **1986**, *Journal of Chemical Education*, 63(7), 617-622.

29. Aebersold, R., Mann, M., **2003**, *Nature*,422,198-207.

30. Loboda, A.V., Krutchinsky, A.N., Bromirski, M., Ens, W., Standing, K.G., **2000**, *Rapid Commun. Mass Spectrom., 72, 552-558.*

31. Roepstorff, P., Fohlman, **1984**, *J. Biomed. Mass Spectrom.*, 11, 601.

32. Papayaannopoulos, I.A., **1995**, *Mass Spectrom. Rev.*, 14, 49.

33. Wattenberg, A., Organ, A.J., Schneider, K., Tyldesley, R., Bordoli, R., Bateman, R.H., 2002, *J Am Soc Mass Spectrom.*, 13, 772-83.

34. Wolfender, J.J., Chu, F., Ball, H., Wolfender, F., Fainzilber, M., Baldwin, M.A., Burlingame, A.L., **1999**, *J. Mass Spectrom.*, 34, 447-454.

35. Hunt, D.F., Yates, J.R.III, Shabanowitz, J., Winston, S., Hauer, C.R., **1986**, *Proc. Natl. Acad. Sci. USA*, 83, 6233-6237.

36. Perkins, D.N., Pappin, D.J.C., Creasy, D.M., Cottrell, J.S., **1999**, *Electrophoresis*, 20, 3551-3567.

37. **http://prospector.ucsf.edu**

38. http://www.matrixscience.com/search_form_select.html

39. Andersson, L., Porath, J., **1986**, *Anal. Biochem.*, 154, 250-254.

40. Holmes, L.D., Schiller, M.R., **1997**, *J. Liq. Chrom. & Rel. Technol.*, 20, 123-142.

41. Ueda, E.K.M., Gout, P.W., Morganti, L., **2003**, *J. Chrom. A*, 988, 1-23.

42. Ren, D., Penner, N.A., Slentz, B.E., Inerowicz, H., Rybalko, M., Regnier, F.E., **2004**, *J. Chrom. A*, 1031, 87-92.

43. Nuhse, T.S., Stensballe, A., Jensen, O.N., Peck, S.C., **2003**, *Mol. Cell. Proteomics*, 2, 1234-1243.

44. Hart, S.R., Waterfield, M.D., Burlingame, A.L., Cramer, R., **2002**, *J. Am. Soc. Mass Spectrom.*, 13, 1042-1051.

45. Campbell, D.G., Morrice, N.A., **2002**, *J. Biomol. Tech.*, 13, 119-130.

46. Salemi, M.R., Qin, Q., Rice, R.H., Lee, Y.M., *Proceedings of the 51st ASMS Conference on Mass Spectrometry and Allied Topics*, Montreal, Canada, June 8-12, 2003; Poster 354.

47. Arnott, D., Gawinowicz, M.A., Grant, R.A., Neubert, T.A., Packman, L.C., Speicher, K.D., Stone, K, Turck, C.W., **2003**, *J. Biomol. Tech.*, 14, 215.

48. Laemmli, U.K., **1970**, *Nature*, 227, 680-685.

49. O'Farrell, P.H., **1975**, *J. Biol. Chem.*, 250, 4007-4021.

50. Gorg, A., Weiss, W., Dunn, M.J., **2004**, *Proteomics*, Nov. 15, p. NA, Epub Ahead of Print.

Chapter 2

Nanoliter Sample Handling, Derivatization Chemistry, and Enrichment of Phosphopeptides by Open Tubular Capillary IMAC.

2.1 Introduction

Phosphorylation and dephosphorylation of serine, threonine or tyrosine is arguably the most commonly encountered post-translational modification of proteins. It has been implicated in wide variety of cellular regulatory mechanisms and transduction pathways[1-5]. Perturbations in the workings of cellular phosphorylation machinery have been implicated in a number of pathological states, including cancer[6-7]. Since concentrations of the kinases and phosphatases which regulate the phosphorylation/dephosphorylation process can vary from high to low, and in addition are often localized to a particular part of an organism or tissue, the resulting stoichiometry of phosphorylation for a cellular extract is in general quite low; a problem compounded by the fact that the site of attachment of the phosphate group on the protein can vary, resulting in a heterogeneous pool of the phosphoprotein. In addition, as discussed in chapter 1, phosphopeptides can be difficult to detect by MS-based technologies due to both ion suppression, and low ionization efficiency of

these peptides in the positive ion mode[8-11]. Therefore, improvements in both sample preparation, and detection sensitivity are critical for future developments of methods for identification and characterization of phosphorylation sites using mass spectrometry.

One method to overcome the foregoing problems is to selectively pre-concentrate any phosphate containing peptides using Immobilized metal-ion affinity chromatography (IMAC). Following binding, clean-up by way of a quasi-selective washing step greatly decreases the proportion of non-phosphorylated and easily ionizable peptides, thereby facilitating detection of phosphopeptides in a much more sensitive manner. Ficarro et al.[12] developed a chemical derivatization step prior to IMAC in an attempt to completely eliminate non-specific binding (i.e., by non-phosphorylated peptides), which presumably occurs through the carboxylate group. By methylating the said group, the authors demonstrated an even greated improvement in the overcoming of ion-suppression effects than with IMAC alone. Both the claims that methylation completely eliminates non-specific binding, and that non-specific binding occurs through the carboxylate group have recently been challenged [13, 14].

In recent years, both instrumental advancements and new sample preparation methodologies have allowed mass spectrometry based technologies to be utilized ever more frequently in the analysis of post-translational modifications of proteins. In our lab, a nanoliter chemistry

station for ultra-sensitive peptide and protein analysis has been developed[15-17]. The salient feature of this approach is the use of fused silica capillaries for handling of nanoliter and sub-nanoliter volumes. Following mixing with MALDI matrix and deposition of microspots on a MALDI target, detection of peptides with a total sample loading of 25000 molecules of the neurotransmitter Substance P has been demonstrated.[18]

In this work, we demonstrate a modified method for nanoliter sample handling using fused silica capillaries which have the metal binding ligand for IMAC attached directly to the silanol groups of the capillary surface. An easy to use hand-held micropipettor-controlled device has been constructed for this purpose. Using this approach we are able to enrich phosphopeptides from fmol-level in-solution and in-gel tryptic digests of the model protein α-casein. Derivatization of carboxylate groups is shown to be an imperfect method to eliminate non-specific binding of peptides to the chelated metal in IMAC; however, application of the negative ion mode in the subsequent MALDI-MS analysis of peptides with esterified carboxylate groups is shown to be an interesting and potentially useful way to selectively ionize phosphate-containing peptides, further increasing the specificity of the method for phosphoprotein analysis.

2.2 Experimental

2.2.1. Materials.

α-Casein and bovine trypsin were obtained from Sigma Aldrich Canada (Oakville, ON). 4-hydroxycinnamic acid (HCCA), acetyl chloride, 2-[N-Morpholino]ethanesulfonic acid, iron(III) chloride, ammonium dihydrogenphosphate, ammonium bicarbonate, hydrochloric acid, trifluoroacetic acid, glacial acetic acid, 3-glycidoxypropyltrimethoxysilane, iminodiacetic acid (IDA) (aka diglycine), sodium hydroxide, and all solvents used were of the highest available purity, and were also obtained from Sigma Aldrich Canada. HCCA was purified by recrystallization from ethanol prior to use. Deionized water was from NANOpure water system (Barnstead/Thermolyne).

2.2.2 Protein Digests and Electrophoresis.

Solution tryptic digests of α-casein were carried out at a concentration of 1 μg/uL with an enzyme to substrate ratio of 1:20 for 24 hours at 37°C. In-gel tryptic digests were performed using the method of Shevchenko *et al.*[19], gel pieces were dehydrated with acetonitrile, dried, and swelled in 10 ng/μl bovine trypsin in 50 mM ammonium bicarbonate; excess trypsin solution was then removed, the gel piece was redissolved in 50 mM ammonium bicarbonate buffer pH 8, and digestion was allow to proceed overnight at 37°C. Peptides were extracted three times with 50% acetonitrile/water/0.2% trifluoroacetic acid, speedvac dried, cleaned by way of commercially available reverse phase C_{18} Zip-Tip^R cartridges (Millipore, Bedford, MA), redissolved in deionized water and stored at -20°C. For IMAC experiments, samples were speedvac dried and reconstituted in 3 μL 0.1 M acetic acid buffer pH 3.5, and loaded directly onto OTC-IMAC. SDS-PAGE was carried out using standard 12%, 0.5 mm thick mini-gels stained with coomasie-blue dye (BioRad, Hercules, CA).

2.2.3. Open Tubular Capillary IMAC Construction, and Phosphopeptide Purification.

A modified method used by Anspach[20-21] was used to attach iminodiacetic acid to the silyl groups on the surface of a fused silica. 4.2 g of NaOH pellets were mixed with 2.8 g of iminodiacetic acid (IDA) in 100 mL of deionized water in a round bottom flask kept on ice; to this solution, 10 mL of 3-glycidoxypropyltrimethoxysilane (GLYMO) were added slowly with stirring. The resulting GLYMO-IDA solution was stirred at 65 °C overnight. About 1 m of a 50 μm i.d., 365 μm o.d. fused silica capillary (Polymicro Technologies, Phoenix, AZ) was activated by flushing it with: NH_3OH:H_2O_2:H_20 (1:1:5) overnight followed by several column volumes of water; with 6 M HCl overnight; and again with several column volumes of water until the pH of the effluent was neutral. The pH of the GLYMO-IDA solution was adjusted to 3.5 with 6 M HCl, and the conditioned capillary was flushed with it overnight at 95°C in a thermostatted oven under nitrogen gas. The capillary was then flushed with a large volume of deionized water, and baked under nitrogen gas at 150 °C.

Open tubular IMAC was constructed from 0.85 mm i.d., 10 cm long glass tubing assembled into a micropipettor-operated device by attaching it to a cut-off pipette tip. Open tubular capillary IMAC (Figure 2.2B) was constructed by cutting 10 mm of the surface modified capillary, and joining it to a cut-off pipette tip through 2 cm of 0.3 mm i.d. Teflon tubing (Supelco, Bellefonte, PA). For IMAC experiments, OTC-IMAC columns were wetted with 50% acetonitrile/0.1% acetic acid, charged with Fe^{3+} by pipetting a solution of 200 mM $FeCl_3$ in 50 mM acetic acid through the capillary; excess iron was washed away with water, and the capillary was then equilibrated with 0.1 M acetic acid buffer (binding buffer).

Phosphopeptides dissolved in binding buffer were loaded by repeated pipetting back and forth through the capillary (no less than twenty times), which was then washed five times with approximately 300 nL of binding buffer and three times with 10% acetonitrile/0.1% acetic acid to further remove non-specifically bound acidic peptides. For derivatized samples, the capillary was washed five times with binding buffer and twice with a solution of 100 mM NaCl containing 25% acetonitrile and 1% acetic acid, in an effort to remove derivatized peptides bound non-specifically through hydrophobic interactions.

Phosphopeptides were eluted with a solution of 200 mM ammonium dihydrogen phosphate (pH 4.5) by pipetting a small plug of approximately 40 nL through the capillary. In order to allow for movement of this solution back and forth through the capillary, and thus improve elution efficiency, a 1 cm long, 300-μm-i.d. Teflon tubing sleeve was attached to the other end of the capillary to prevent loss of analyte. An equal volume of 10 mg/mL HCCA matrix containing 1% HCl was mixed with the eluting solution inside the capillary and connecting Teflon tubing. The resulting solution was then spotted onto a thin first layer of HCCA matrix according to a modified version of the two-layer method[22]. One percent HCl was originally used in order to acidify pH 9 ammonium phosphate to prevent dissolution of the first layer of HCCA matrix, and was observed to have no detrimental effect on signal intensities. It was fortuitous to discover later that including this additive with the pH 4.5 ammonium phosphate made for much easier spotting from the capillary IMAC, perhaps due to the high viscosity of HCl. The spots resulting from this methodology were then washed with deionized water and analyzed by mass spectrometry.

For experiments involving derivatization of carboxylate groups without modifying the phosphoryl residue, the method of Ficarro *et. al.*[12] was used. Briefly, 1 M methanolic HCl was prepared by addition of 200 μL acetyl chloride, with mixing, to 2.5 ml of absolute methanol through a septum with a gas-tight syringe. One μl of this solution was added to speedvac-dried phosphopeptide mixtures and the derivatization was allowed to proceed for 2 h at room temperature. Samples were then speedvac dried once more, reconstituted in 10 μL 0.1% trifluoroacetic acid and purified by reverse phase C_{18} Zip-Tip^R cartridges, dried, and redissolved in 1 μL HCCA matrix for direct analysis, or 3 μL binding buffer for loading onto IMAC columns.

For the sensitivity study, open-tubular IMAC and commercially available IMAC were applied using the same binding and washing conditions as those described for OTC-IMAC above. Phosphopeptides were bound from 10 μL of binding buffer, followed by washing thrice with the same buffer and thrice with a solution of 10% acetonitrile/0.1% acetic acid. Elution from OT-IMAC was achieved using 5 μL 200 mM pH 4.5 ammonium phosphate solution; Elution from Millipore-IMAC^{MC} was with 5 μL 0.3% NH_4OH, according to manufacturer's protocol. Spotting by the two-layer method for HCCA and the dried droplet method with DHB.

2.2.4. Mass Spectrometry.

MALDI was performed on a Voyager Elite reflectron time-of-flight mass spectrometer (Framingham, MA) equipped with a nitrogen laser (λ = 337 nm) by collecting signals averaged from 200-400 laser shots. Spectra were calibrated externally using monoisotopic ions of a pepcal standard, or internally using trypsin autolysis peaks. MALDI spectra were analyzed using the Igor Pro Software package (WaveMetrics, Lake Oswego, OR).

2.3. Results and Discussion

2.3.1. Open Tubular Capillary IMAC (OTC-IMAC): Construction and Performance.

As mentioned above, trace analysis of peptides in our laboratory was traditionally done using the nanolitre chemistry station, shown in Fig. 2.1. In short, this approach utilizes a XYZ stage and a microscope for viewing a capillary to be used for microspot sample handling. The movement of liquids in and out of the capillary is done using a piston driven syringe. This main drawback of this method is the difficulty associated with manual manipulation of ultra small volumes, and concomitant poor reproducibility. Although this approach is excellent for testing the fundamental detection limits in MALDI-MS, owing to the picoliter volumes used for elution, we wanted to develop a more robust and user friendly method, whilst still keeping the inherent low volume advantage inherent in the use of fused silica capillaries for sample analysis and preconcentration.

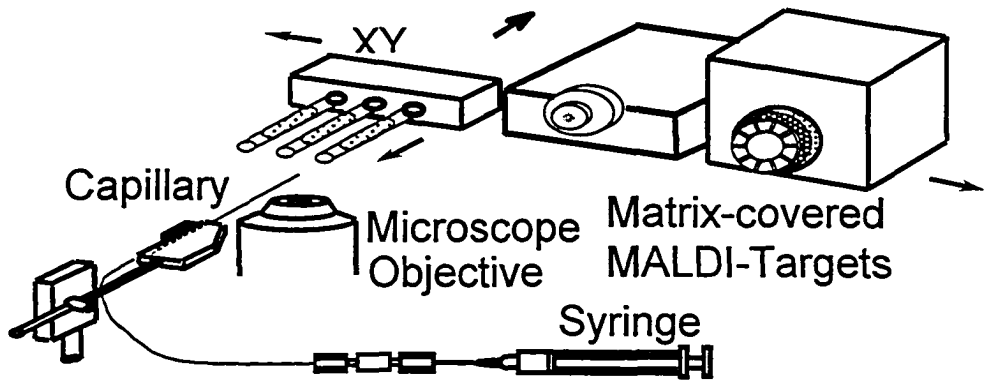

Figure 2.1. The nanolitre chemistry station.

To meet the goal described above, we constructed an affinity capillary by attaching the diglycine ligand which forms a phosphate-binding coordination complex with various metals directly to the silanol groups on the inner surface of the fused silica. The surface-modified open tubular capillary IMAC (OTC-IMAC) was assembled into a hand-held micropipettor-controlled device. Initial development of this device was done on a large scale using open tubular IMAC in 10 cm long and 0.85 mm i.d. glass tubing; the performance and application of this device was recently been published[23]. We also assembled the glass tubing open tubular IMAC (OT-IMAC) synthesized by Dr. Huaizhi Liu, a postdoctoral fellow of our lab, into a micropipettor-operated device for convenience. Both OT-IMAC and OTC-IMAC are shown in Fig. 2.2.

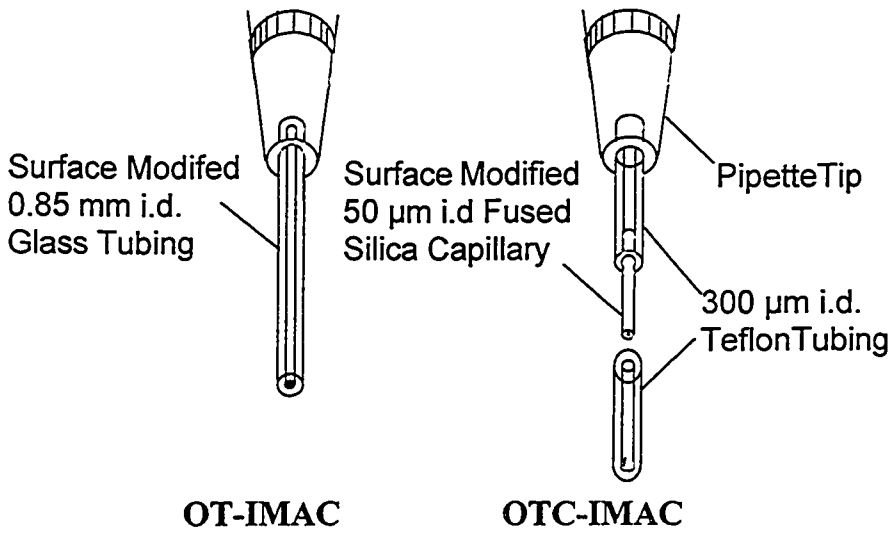

Figure 2.2. The open tubular IMAC (OT-IMAC) and open tubular capillary IMAC (OTC-IMAC) presented in this work.

As mentioned above, by using the nanolitre chemistry station, it is possible to handle picolitre volumes in a fused silica capillary. The herein described OTC-IMAC device allows for handling of volumes in the 25-100 nL range, with the use of 50 μm inner diameter capillaries. This range could be decreased further by using thinner capillaries; however movement of liquids through capillaries thinner than 50 μm then becomes very difficult and time consuming. Hence, the OTC-IMAC device is a compromise between robustness and usability on the one hand, and a relatively low sample volumes on the other. Typical spots obtained using the OTC-IMAC device are shown in figure 2.3C. On average, the resulting MALDI spots are 0.5 ± 0.04 mm in diameter, although this depends largely on the skill of the operator. Fig. 2.3A and B shows OTC-IMAC in action, as eluate is expelled onto a MALDI target. Considering a typical diameter of 2.5 mm for regular MALDI spots from a 1 μL volume, OTC-IMAC can potentially improve signals by a factor of approximately 25, assuming linearity of signal versus concentration. This is due to the fact that our analyte is concentrated into an area one twenty-fifth the size of what it would be under normal conditions.

$$V_{OTC\ SPOT} / V_{REG.\ SPOT} = \pi R^2 / \pi R^2$$

$$= \pi (1.25\ mm)^2 / \pi (0.25\ mm)^2$$

$$\approx 25$$

Hence, the OTC-IMAC device under discussion does not produce true 'micro-spots' as defined by micro-spot pioneers in our lab[15-18]; however, it still holds a definite advantage over traditional spotting technique. In addition, a 2.5 mm diameter spot would come from, typically, a 1 μL volume solution; however elution of peptides from any macro IMAC device becomes impractical with less than about 4 μL of eluant, owing to the bed volumes typically present in such devices, and the same is true for the OT-IMAC constructed in this lab. Hence, the calculated 25-fold improvement in sample concentration would represent a minimum value.

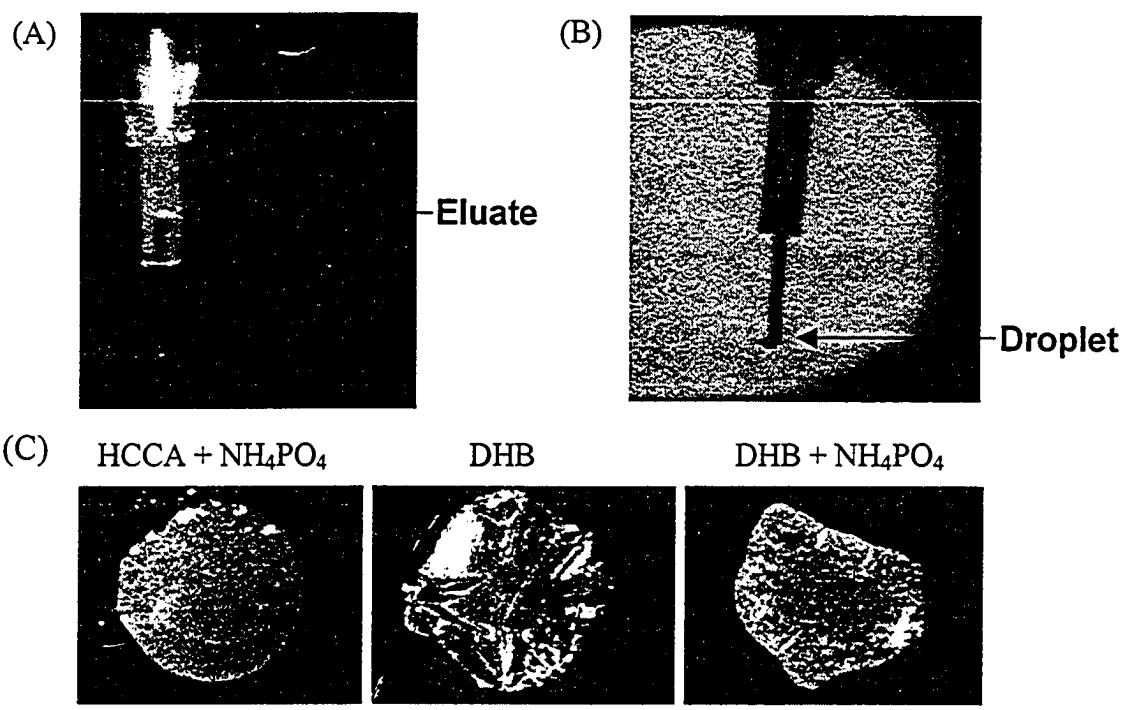

Figure 2.3. Microscope images of the microspot deposition technique used with OTC-IMAC. A: OTC-IMAC filled with a sample plug, positioned above a steel MALDI target; B: Expulsion of the sample plug via a micropipettor onto a MALDI target; C: Images of the spots resulting from OTC-IMAC deposition.

2.3.2. Influence of Chelated Metal on OTC-IMAC Performance

As mentioned in Chapter 1, several metals have been used for peptide purification and enrichment in IMAC. Most commonly, Fe^{3+} or Ga^{3+} have been used for this purpose[24-28] No study exists comparing the influence of different metal-binding ligands on the efficiency in IMAC of phosphopeptides, although Ren *et. al.*[29] found that different commercially available sorbents have disparate selectivities when used for histidine-tagged protein purification on Cu(II) IMAC. Although a study comparing the influence of both the different metals, metal-binding ligands, and the embedding matrices to which the ligands are attached would be interesting—since there is no common consensus as to which product on the market is best—such a comprehensive undertaking has never been carried out.

In the herein presented work, the metal chelating ligand diglycine has been attached directly to the silanol groups of fused silica—and analogously to the silanol groups of glass in the work published by Liu H.; there are no frits or embedding/supporting matrices such as gels, polymers, or glue, and in addition the surface chemistry of fused silica or glass is more amenable to control on account of the homogeneity found in these surfaces, as opposed to say, porous chromatographic beads.[23] Thus, we would expect our approach to result in a more efficient and reliable phosphopeptide capture device for IMAC enrichment.

We tested the performance of OTC-IMAC charged with several different metals: Al(III), Cu(II), Ga(III), and Fe(III) present as the chloride salts. Phosphopeptides were bound under identical buffer conditions. Following equilibration with binding buffer and capture of phosphopeptides from a tryptic digest of the model phosphoprotein α-casein, the same number of washing steps were performed. We found Fe(III) to be by far the best metal for our method. These results are shown in figure 2.4. Posewitz *et al.*[26] found Ga^{3+} to be most efficient and selective for phosphopeptide enrichment with Poros-IDA beads packed into a gel-loader tip, as compared to Fe(III), Al(III) and Zr(IV). Others have reported increased selectivity in Al(III) IMAC over Fe(III) IMAC, with nitrilotetraaceticacid-ligand based columns.[30-31] The superiority of Zr(IV) in IMAC has also been observed combined with a "new chelate" developed by Merck Frost[32]. The upshot from such reports is that all methods for doing IMAC differ in their optimum capture conditions, depending on both the type of ligand used (for example, IDA vs. NTA), and the metal which is immobilized on a surface by way of the said ligand; in addition, different support matrices used to immobilize the ligand itself, are likely to play a role, especially in regards to non-specific background binding.

From our results, it appears that Cu(III) is the least efficient at capturing phosphopeptides, as evidence by the presence of only four of these upon elution. Ga(III) and Al(II) appear to be quite good, as evidenced

by the appearance of several new peptides compared to the pre-IMAC control; however the concurrent absence of some peptides observed with Fe(III) enrichment, and the presence of a large number of non-specifically bound peptides, led us to favour enrichment with Fe(III). Perhaps with more extensive washing to eliminate non-specific binding the performance of Ga and Al could possibly be improved a little; however, since we detected the greatest number of phosphopeptides with Fe(III), because the number of washing steps used is already somewhat tedious and time consuming, and because the number of non-specifically bound peptides was also the lowest with this metal, it was chosen for all further experiments.

Although for all practical purposes the above conclusion is the right one, strictly speaking it is quite conceivable that the different metals used here have completely different binding/metal-affinity characteristics; that is, they may each and every one have different optimum conditions for binding phosphopetides (for example, pH or type of binding buffer, preference for organic solvents, ionic strength, etc.). Hence, to unequivocally conclude that Fe(III) is absolutely the best metal for purposes of phosphopeptide capture with fused silica-IDA, one would have to perform any number tests to optimize all conditions for each metal separately.

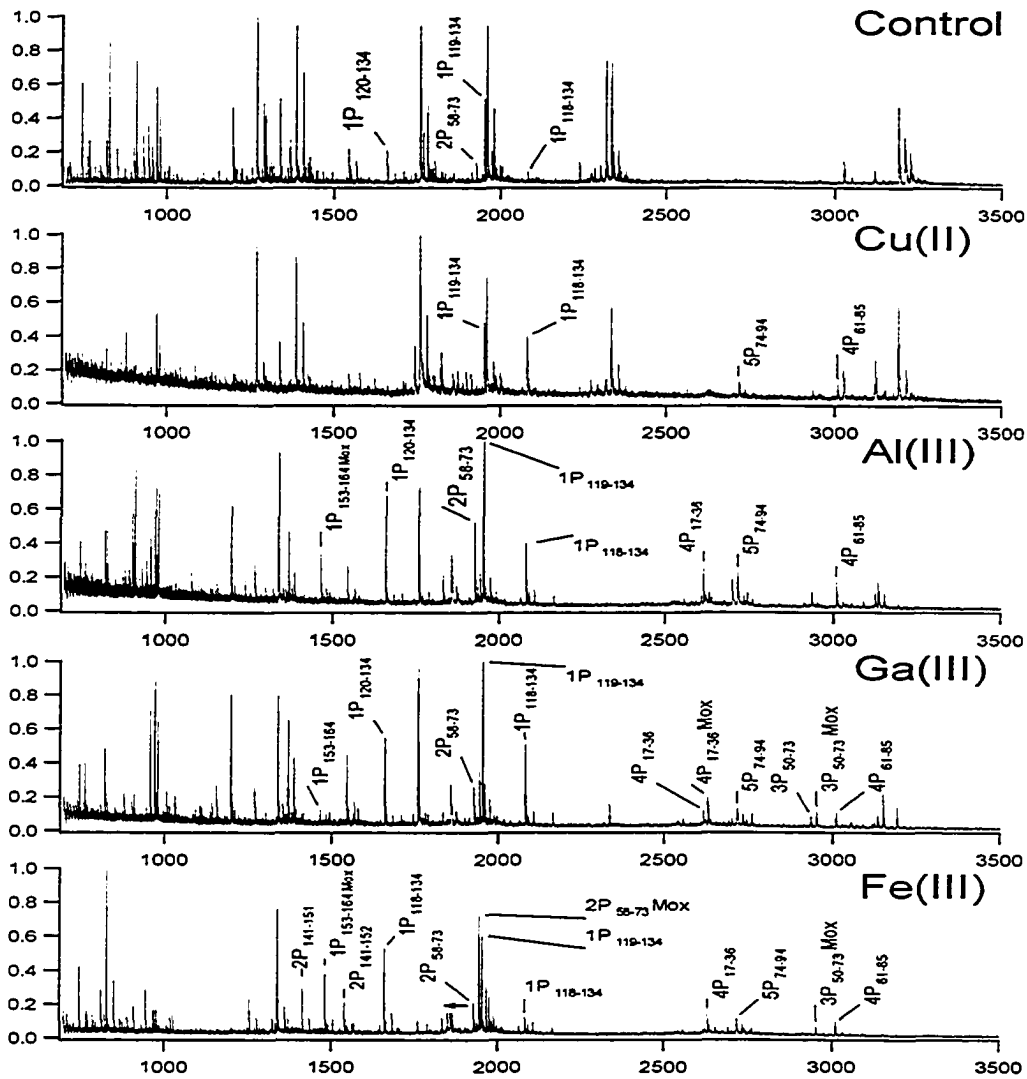

Figure 2.4. Metal-Dependent Performance of OTC-IMAC. A 1 pmol tryptic digest of α-casein was used for these experiments. Binding of phosphopeptides was from a 0.1 M acetic acid solution; following binding, 5 washing steps with binding buffer, and 5 washing steps with 0.1 M acetic acid/15% acetonitrile were performed. Elution was by 0.2 M $NH_4H_2PO_4$, and spotting for MALDI utilized DHB matrix .

2.3.3. Not All Phosphopeptides are Equal

In a review discussing the practical aspects of phosphoprotein characterization, it has been pointed out that IMAC enrichment has been successfully applied for enrichment of "model proteins", but remains notoriously unreliable for "real life" unknown proteins.[33] Other studies

involving *en masse* phosphopeptide enrichment from entire proteomes, have pointed to the fact that such approaches lead to the detection of a disproportionately high number of multiply phosphorylated peptides[12,13]—leading to an unrealistic under-representation of peptides with only a single phosphate. Such observations have been rationalized in terms of the expected higher affinity of IMAC for peptides bearing multiple phosphate residues.

In our method development for OTC-IMAC, it was noticed that the solution remaining after application of IMAC—what is termed the 'flow-through'—still contained some phosphopeptides. Furthermore, the wash solution used to minimize non-specific binding also contained the very same phosphopeptides, along with a number of other, non-phosphorylated peptides. This data is shown in Figure 2.5. Comparing panel A and D, it is obvious that the OTC-IMAC device is binding phosphopeptides just fine, and that most non-phosphorylated peptides are being eliminated from final detection. Overall, IMAC is very effective at enriching phosphopeptides; however, as is evident in panels B and C, the two singly phosphorylated peptides do not undergo an ideal 'affinity capture' process; instead, these peptides appear to be present in equilibrium with the immobilized metal, with only a portion of the peptide being present in the metal-bound phase. This would explain the observation shown in panel C, namely that these peptides are much easier to wash away than some of the others. In a poster

presented at the 2003 American Society for Mass Spectrometry Conference, Voisin *et. al.*[34] found that IMAC preferentially binds some phosphopeptides, while leaving others in solution. It is also quite conceivable that some peptides simply do not bind IMAC because of unusual folding, making access to the phosphated group difficult for the chelated metal; however, at present, there are no strategies to either study or overcome such limitations.

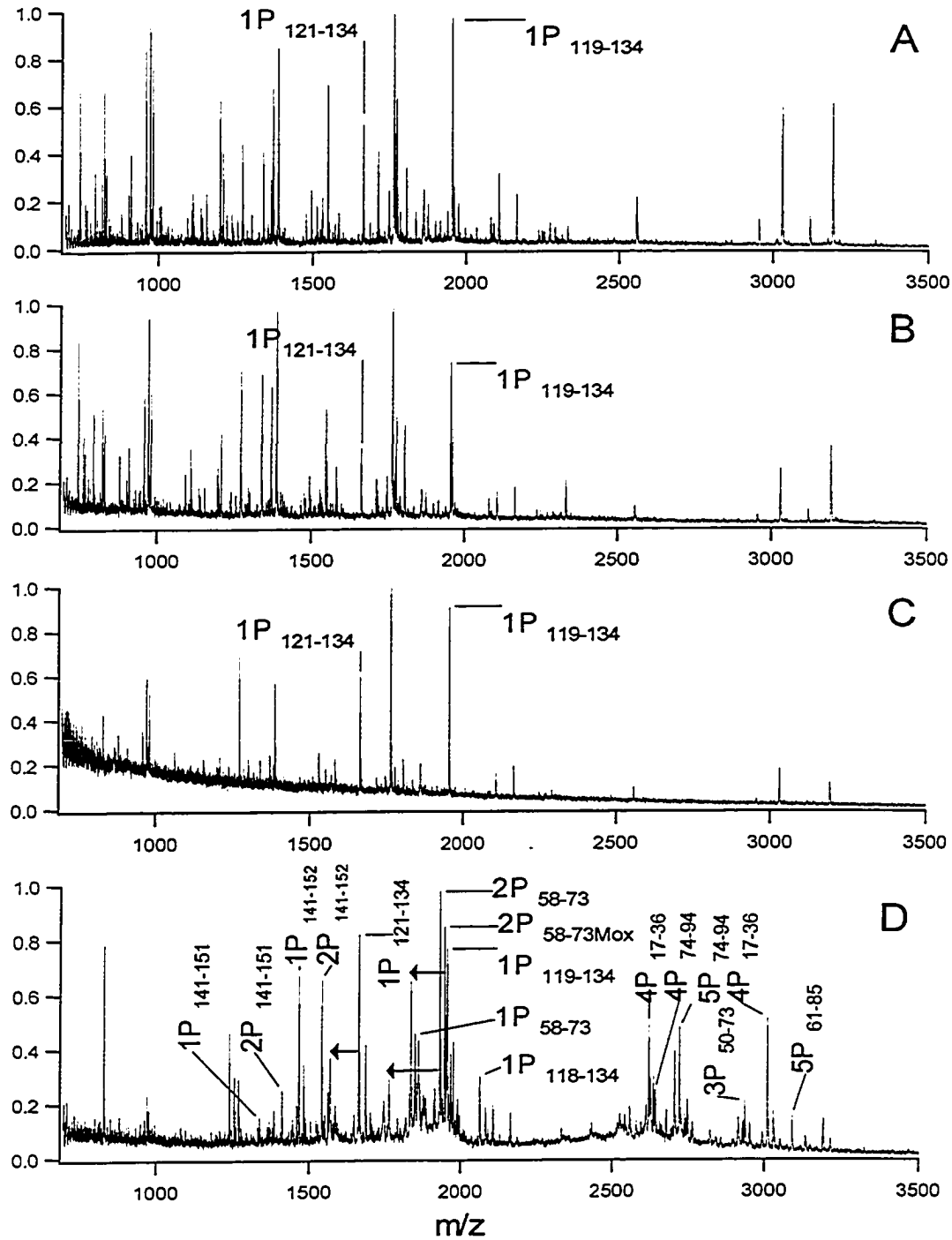

Figure 2.5. MALDI-MS analysis of the flow-through and wash obtained with OTC-IMAC enrichment of 5 pmol tryptically digested α-casein. A: No IMAC enrichment. B: Flow-through solution after phosphopeptides were fished out with OTC-IMAC. C: Wash solution meant to remove non-specifically bound peptides. D: Eluate from OTC-IMAC after the washing step.

2.3.4. Influence of Instrumental Setup and Matrix in MALDI-MS: Positive Versus Negative-Ion Mode MALDI-MS Without IMAC enrichment

In some of the earlier published work on phosphorylation analysis, several authors suggest using linear as opposed to reflectron-mode MALDI-MS for more sensitive phosphopeptide detection. This is due to the fact that being somewhat labile, the phosphate group has been known to undergo metastable decay, giving rise to –95 Da fragments from the original peptide. It was also suggested that negative-ion mode MALDI-MS could be a way to enhance the relative ionization efficiency of phosphate carrying peptides, and that assigning peaks with increased relative intensities in the negative mode could be a way to identify potential phosphopeptide candidates[8-11] without use of a preconcentration step. In general, we found these methods to be very unreliable and not of great use. Figures 2.6 and 2.7 show the results for 1 pmol α-casein without IMAC enrichment, analyzed in positive/negative and in linear/reflectron modes, with HCCA and DHB MALDI matrices respectively. In looking at these figures, the only salient feature is that the relative peak intensities sometimes do increase (usually they don't change very much), but it really depends relative to what. The intensities of the non-phosphorylated peaks change between the analyses in positive and negative ion modes, but they change very differently for many of the peptides in this very complex

mixture, so actually saying that from this kind of analysis one could pick out phosphopeptide candidates is not very convincing. Furthermore, our process of acquiring MALDI-MS spectra involves averaging many 'hot-spots' where higher signal intensities are found. These spots arise from the process of crystallization of matrix and sample, and don't all contain the same ratios of all the peaks. Averaging too few spectra could easily lead to erroneous conclusions. Also of note is the fact that any talk of 'hotspots' only applies to DHB matrix, since with the two-layer used for HCCA crystallization in our laboratory produces uniform crystallization, and hence there is no need to hunt for 'hotspots'; however, with some minor differences between the total number of phosphopeptides detected, a similar trend was observed with both HCCA and DHB matrices, although DHB does show a little better ionization selectivity in the negative-ion mode than HCCA.

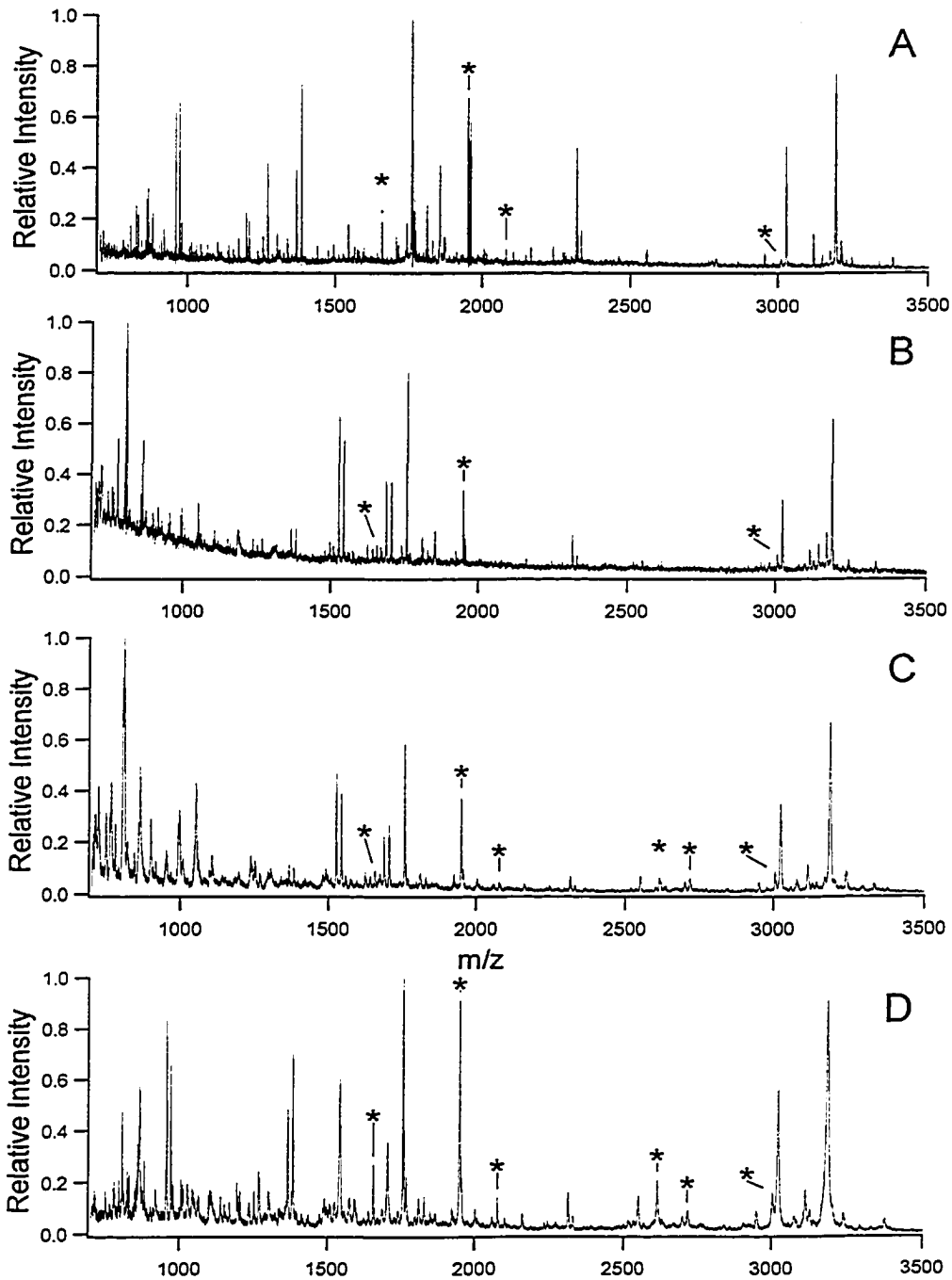

Figure 2.6. Comparison of tryptically digested α-casein in reflectron and linear mode MALDI-MS in the positive and negative ion modes with HCCA as ionization matrix. 1 pmol of tryptically digested α-casein was used for this analysis. A: Reflectron/positive mode. B: Reflectron/negative mode; C: Linear/positive mode; D: Linear/negative mode.

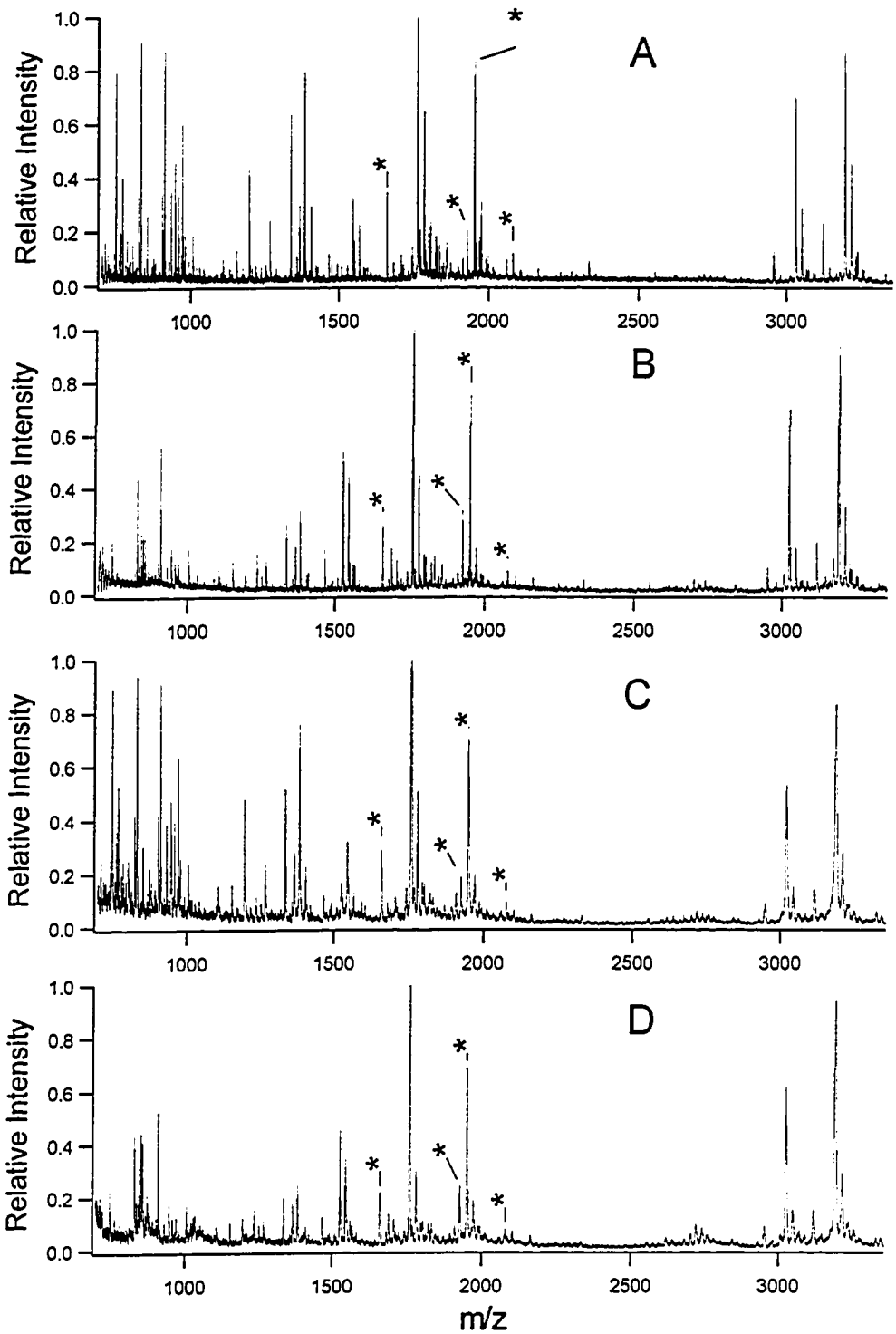

Figure 2.7. Comparison of tryptically digested α-casein in reflectron and linear mode MALDI-MS in the positive and negative ion modes with DHB as ionization matrix. 1 pmol of tryptically digested α-casein was used for this analysis. A: Reflectron/positive mode. B: Reflectron/negative mode; C: Linear/positive mode; D: Linear/negative mode.

2.3.5. Positive Versus Negative Mode MALDI-MS Analysis of OTC-IMAC Enriched Phosphopeptides: Reflectron and Linear TOF MALDI-MS.

We also wanted to test the foregoing observations following IMAC enrichment, in order to have a complete picture of instrumental setups available for this kind of analysis. 1 pmol tryptically digested α-casein was enriched with OTC-IMAC, the different MALDI-MS modes of analysis were then systematically investigated: we compared the spectra of the IMAC enriched α-casein in linear and reflectron MALDI-MS in positive and negative ionization mode; in addition, as in the foregoing section, the same comparisons were made with both HCCA and DHB as the ionization matrix.

OTC-IMAC enriched trypsinized α-casein analyzed in positive-ion mode with HCCA matrix is shown in Fig. 2.8; the figure compares the reflectron and linear modes. It is apparent that OTC-IMAC allows us to preconcentrate phosphopeptides and overcome ion-suppression effects, as evidenced by the greater number of phosphopeptides detected after enrichment than without it (compare Fig. 2.8A versus Fig 2.6A). It is also evident that with HCCA as matrix, losses of the phosphate group are occurring in flight, as shown by the satellite ions (-95 Da) marked with arrows in Fig. 2.8. These peaks can be useful, as they can be an early diagnostic for the presence of the phosphate group, although the intact

peptide with the phosphate group still attached suffers a decrease in sensitivity because of this phenomenon.

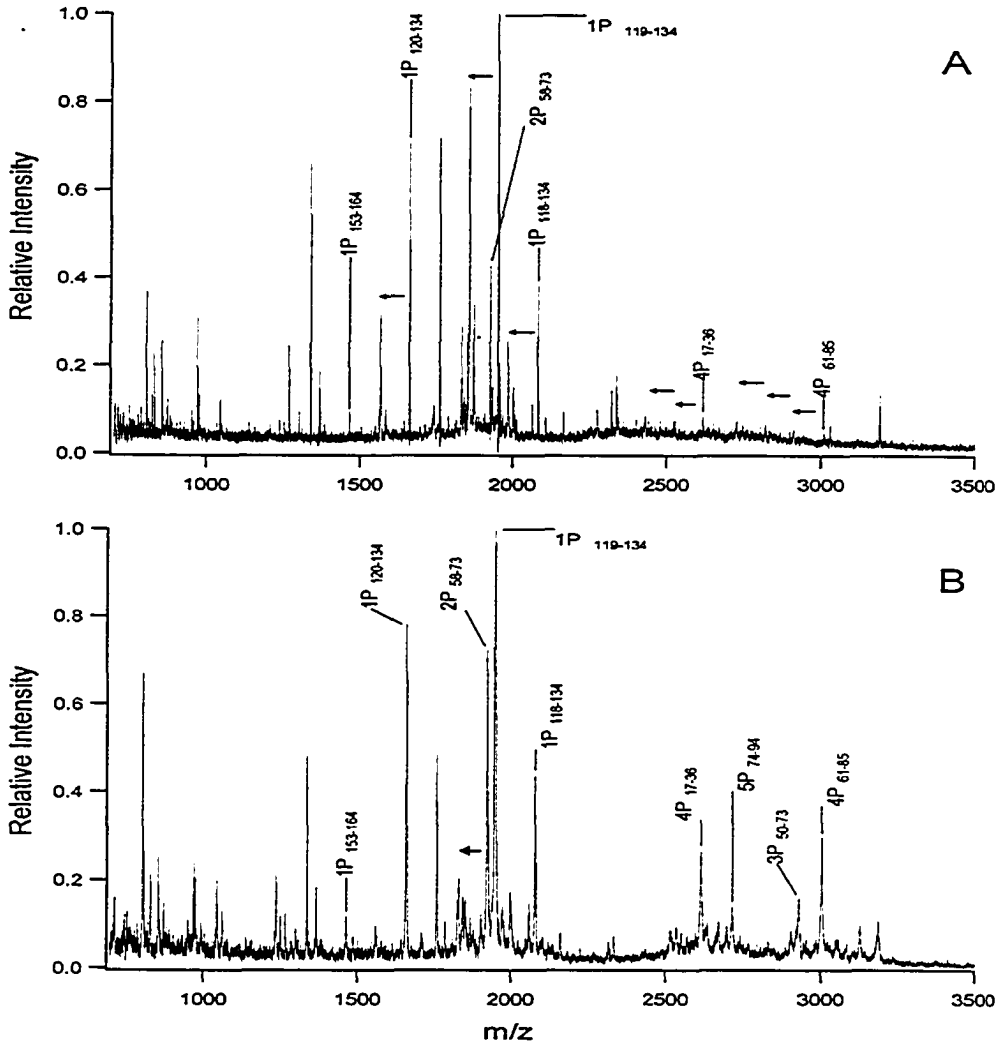

Figure 2.8. Comparison of tryptically digested α-casein in reflectron and linear mode MALDI-MS in the positive ion mode with HCCA as ionization matrix. A: 1 pmol OTC-IMAC enriched α-casein; reflectron mode. B: 1 pmol OTC-IMAC enriched α-casein; linear mode.

The reason for this loss of the phosphate group is the ion higher energy transfer from the HCCA matrix. The reflectron allows detection of ions which have fragmented in flight, which does not occur in linear mode as the ions travel together, as shown in panels A and B. Hence there is a

concomitant increase in sensitivity in the linear mode due to the ions traveling to the detector fully intact. This effect on sensitivity seems to be most pronounced for the multiply phosphorylated peptides in the high-mass range, since they are able to lose more than one phosphate group. The obvious drawback of the linear mode is of course the lack of isotopic resolution, and hence a much decreased confidence in identification of peptides.

Next we made the same comparison as above in the negative ion mode. As shown in Fig. 2.9, a very similar trend is observed in negative mode as was seen in the positive mode (Fig. 2.8), including the increase in sensitivity in the linear mode. One notable difference however is seen when Fig. 2.8A and Fig. 2.9A are compared. It appears that following IMAC enrichment, the effect of increased relative peak intensities is negative-ion mode is a little easier to notice than without enrichment. However, the effect here is hardly pronounced enough to be made into a criterion for phosphopeptide identification, and appears to become more significant as the complex mixture is simplified through affinity chromatography.

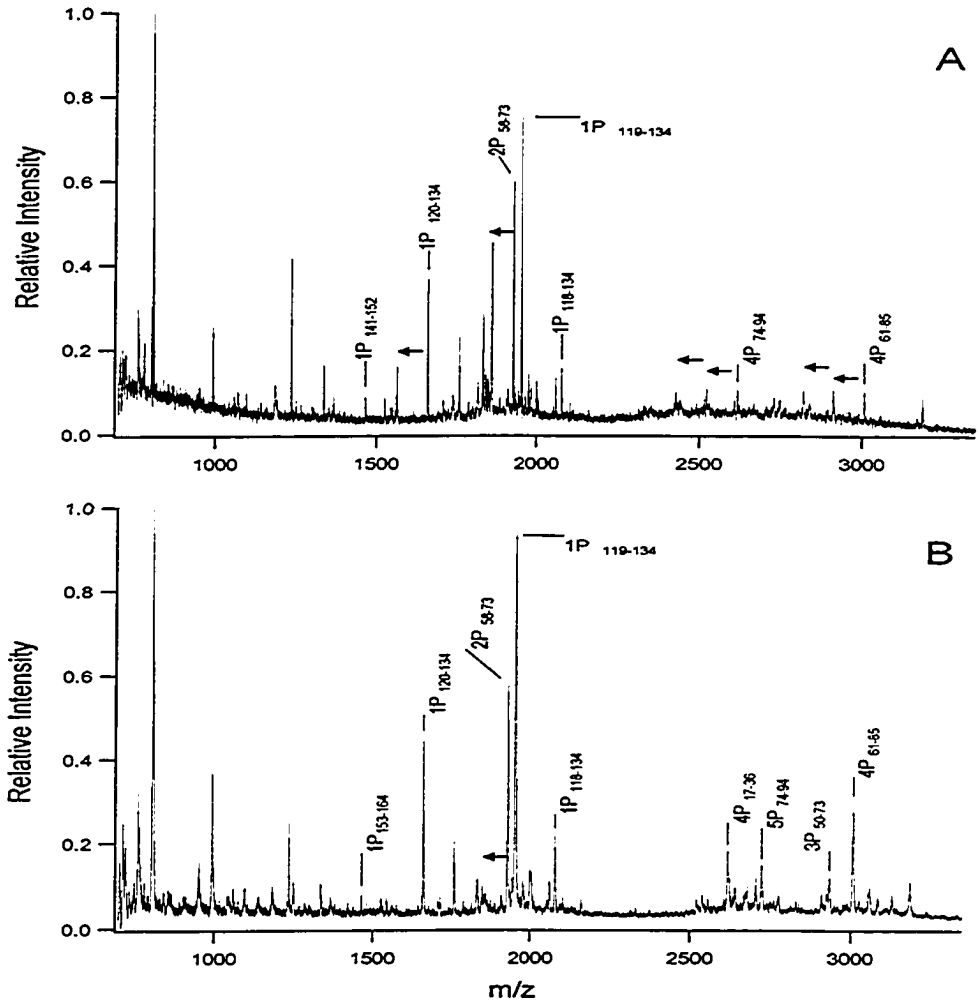

Figure 2.9. Comparison of tryptically digested α-casein in reflectron and linear mode MALDI-MS in the negative-ion mode with HCCA as ionization matrix. A: 1 pmol OTC-IMAC enriched α-casein; reflectron mode; B: 1 pmol OTC-IMAC enriched α-casein; linear mode.

Figures 2.10 and 2.11 show the same experiments as the foregoing, with DHB as the ionization matrix. The main difference between DHB and HCCA is DHB's property of being a 'soft' matrix, whereas HCCA is termed a 'hard' matrix[24]. A much lower laser energy is required to ionize peptides using HCCA than DHB, which suggests that this particular matrix transfers energy to the analyte in a more efficient fashion; more energy is

imparted to the molecules, hence giving less stable ions more likely to fall apart in flight. This is what is believed to cause the increase in metastable decay of post-translationally modified molecules such as phosphopeptides when HCCA is used, as discussed above. The positive mode reflectron versus linear mode MALDI-MS analysis of OTC-IMAC enriched α-casein analyzed with DHB matrix is shown in Fig.2.10. Fig. 2.11 shows the same data in negative-ion mode. The salient feature in comparing Fig.2.10 (DHB) with Fig.2.8 (HCCA) is the increase in sensitivity with DHB as the matrix, without the sacrifice in resolution required to achieve such an increase by applying the linear mode when HCCA was used. Also notable is the presence of a higher number of phosphopeptides detected with DHB. When we compare Fig.2.10 and 2.11 directly, we can see that the relative peak intensities in negative-ion mode do increase more with DHB than they did with HCCA; however this is true only for IMAC-enriched peptides, and was not observed to be of any significance otherwise. In any event, we conclude that DHB is indeed the best matrix to use for analysis of phosphopeptides on account of its greater sensitivity and usefulness in the reflectron mode, which also gives superior resolution.

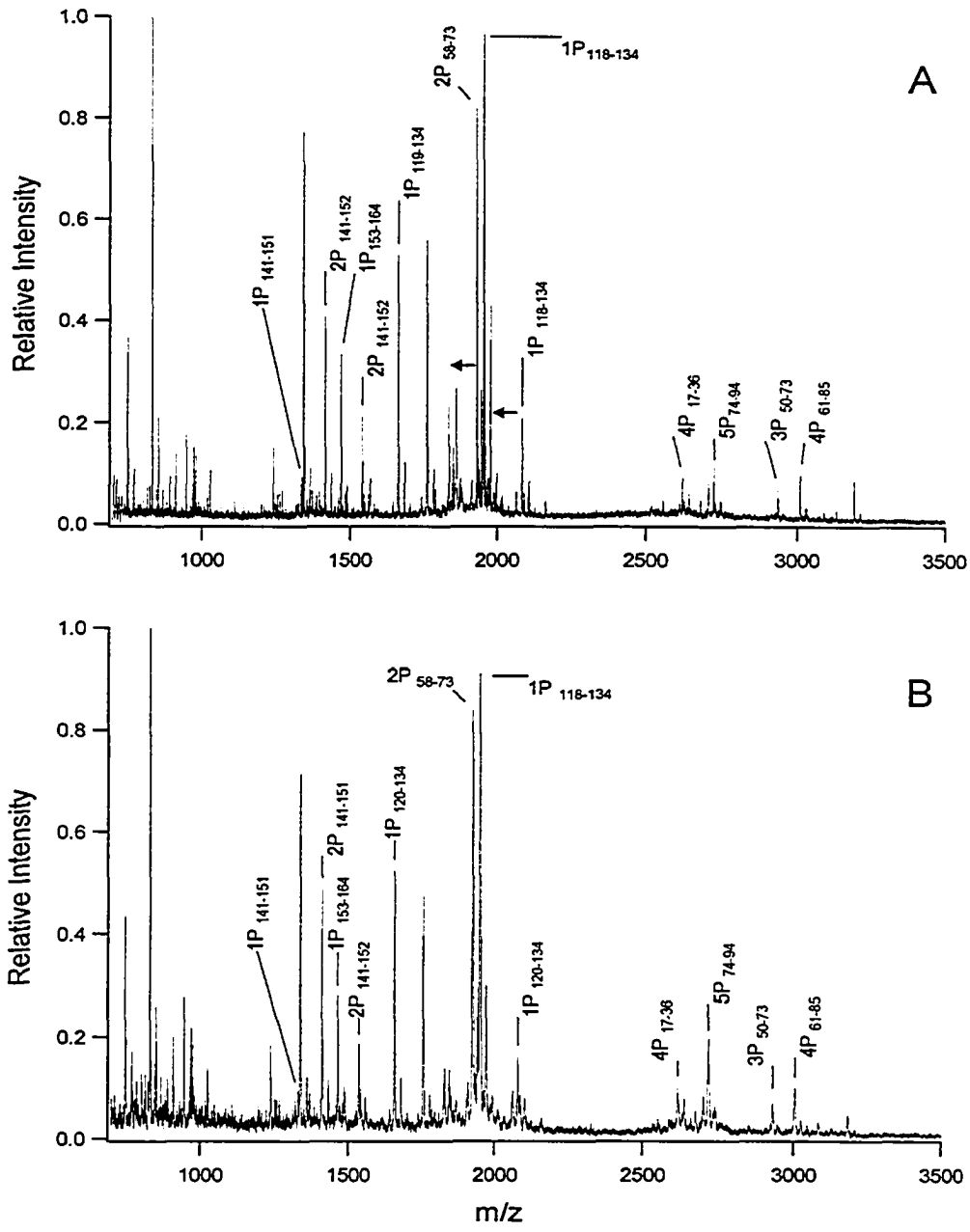

Figure 2.10. Comparison of tryptically digested α-casein in reflectron and linear mode MALDI-MS in the positive-ion mode with DHB as ionization matrix. A: 1 pmol OTC-IMAC enriched α-casein; reflectron mode; B: 1 pmol OTC-IMAC enriched α-casein; linear mode.

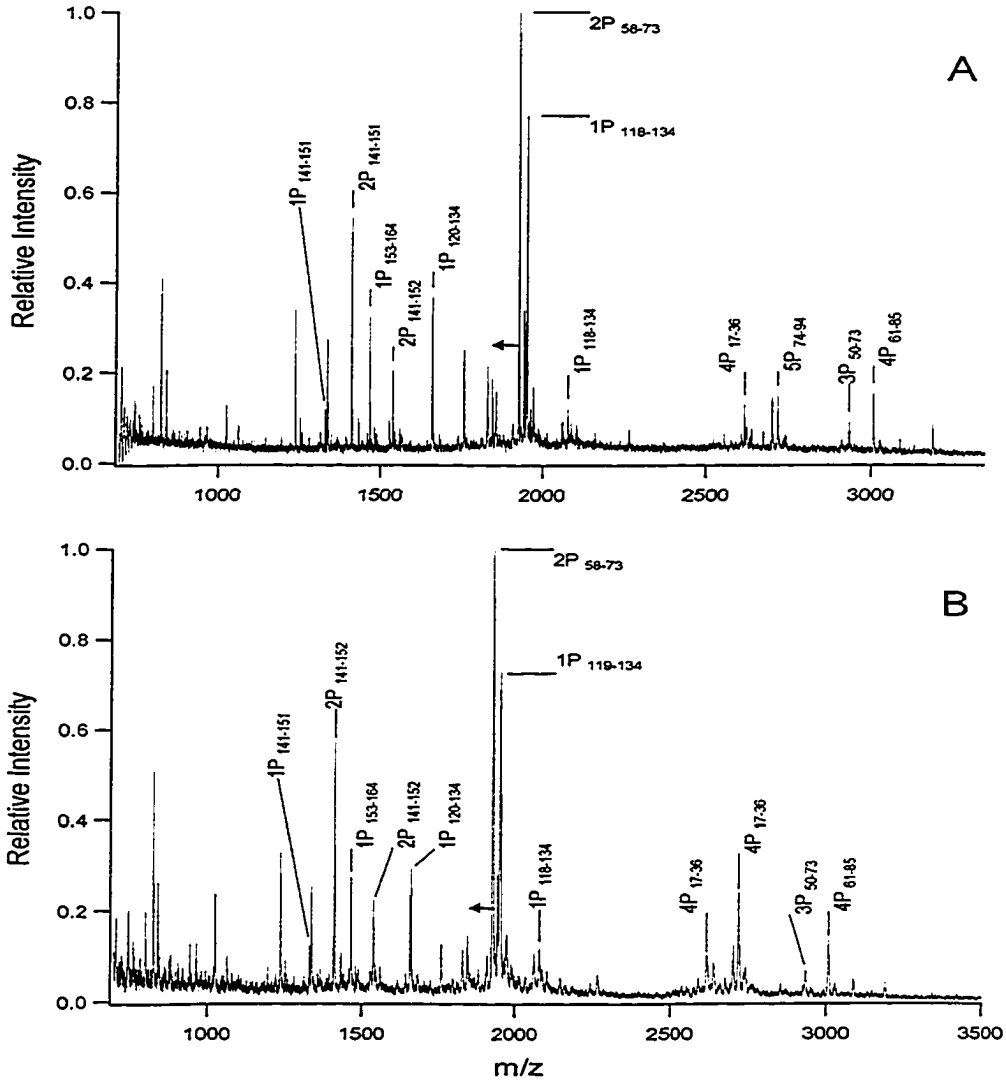

Figure 2.11. Comparison of tryptically digested α-casein pre-IMAC and post-IMAC in reflectron and linear mode MALDI-MS in the negative ion mode with DHB as ionization matrix. A: 1 pmol OTC-IMAC enriched α-casein; reflectron mode; B: 1 pmol OTC-IMAC enriched α-casein; linear mode.

2.3.6. Detection Sensitivity of OTC-IMAC and Comparison to Commonly Used Commercial IMAC Product.

The idea behind constructing the herein presented IMAC devices was to improve on the performance of commercially available packed pipettor tip devices, which workers at the Alberta Cancer Board Proteomics

Resource Facility have had limited success with for real world sample applications. As mentioned before, our devices contain no frits, no embedding or supporting matrices or non-IMAC filler materials; no gel, polymer, or glue is present, only silica derivatized with the diglycine functionality which binds the phosphopeptide-binding metal directly.

We set out to compare the sensitivity of our devices versus that obtainable with the commercial (and very expensive!) product from Millipore by analyzing a 1 pmol tryptic in-solution digest of α-casein. We gauged the sensitivity for IMAC enrichment using both HCCA and DHB matrices. The data for this analysis using HCCA matrix is shown in Fig. 2.12; that obtained with DHB as the MALDI ionization matrix is shown in Fig.2.13. As evident from the two figures, DHB is the more sensitive matrix—as was expected. The sensitivity trend observed is clearly OTC-IMAC > OT-IMAC > MilliPore IMAC. The small elution volume advantage of the OTC-IMAC device is apparent from both the higher total number of phosphopeptides detected, and their intensities. Figure 2.14 shows the spectrum obtained with OTC-IMAC enrichment from a 100 fmol α-casein digest. All the peaks observed with 1 pmol are still detectable. As shown in panel B of the same figure, OTC-IMAC enrichment is still feasible at 20 fmol; however, the total number of phosphopeptides observed

naturally decreases at this very low concentration.

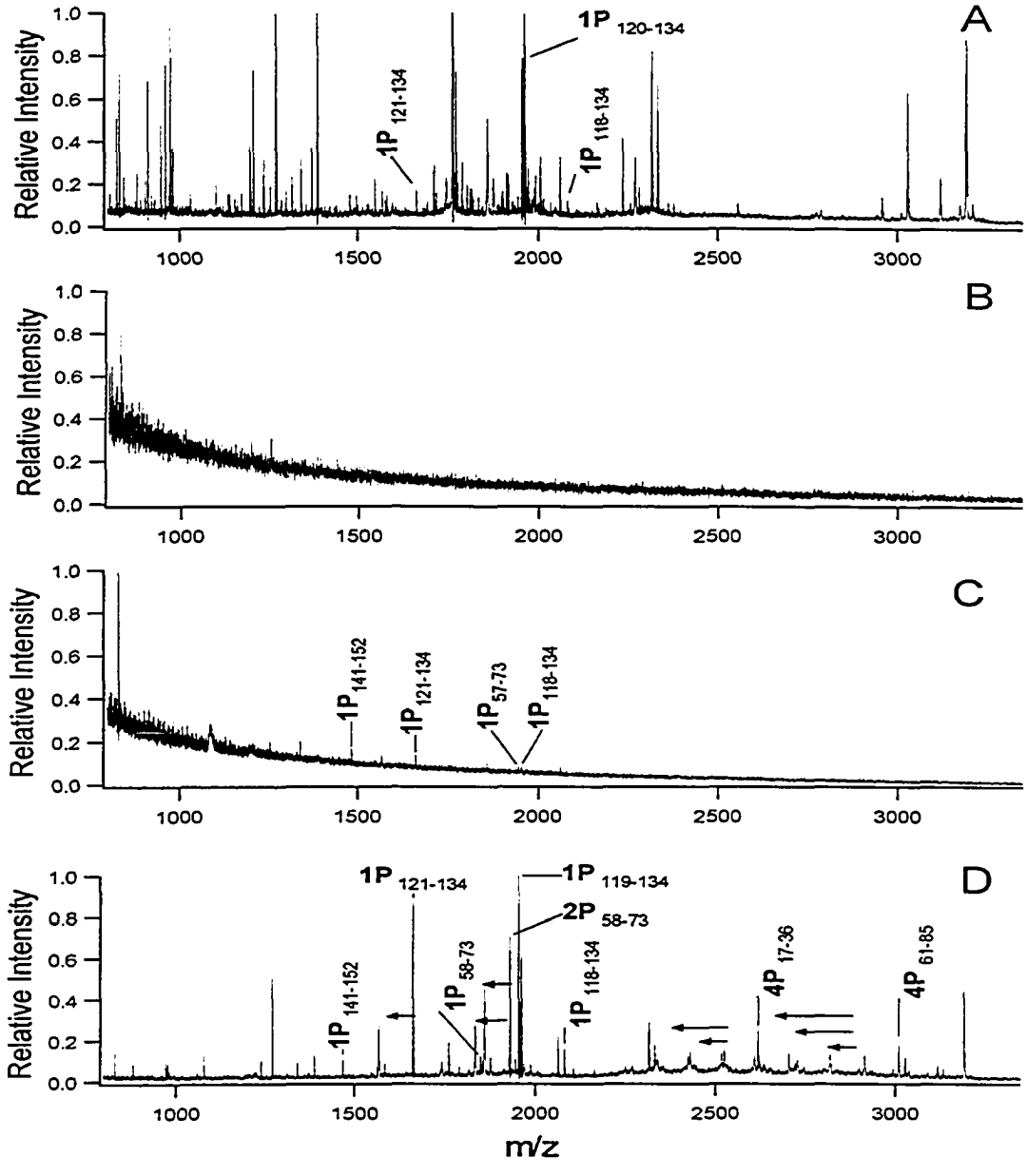

Figure 2.12. Detection sensitivity comparison between the commercially available packed pipettor tip IMAC from Millipore, OT-IMAC, and OTC-IMAC gauged by MALDI-MS with HCCA matrix. A 1 pmol tryptic digest of α-casein was used for this purpose. A: digest without IMAC enrichment; B: Millipore-IMAC enriched digest; C: OT-IMAC enriched digest; D: OTC-IMAC enriched digest.

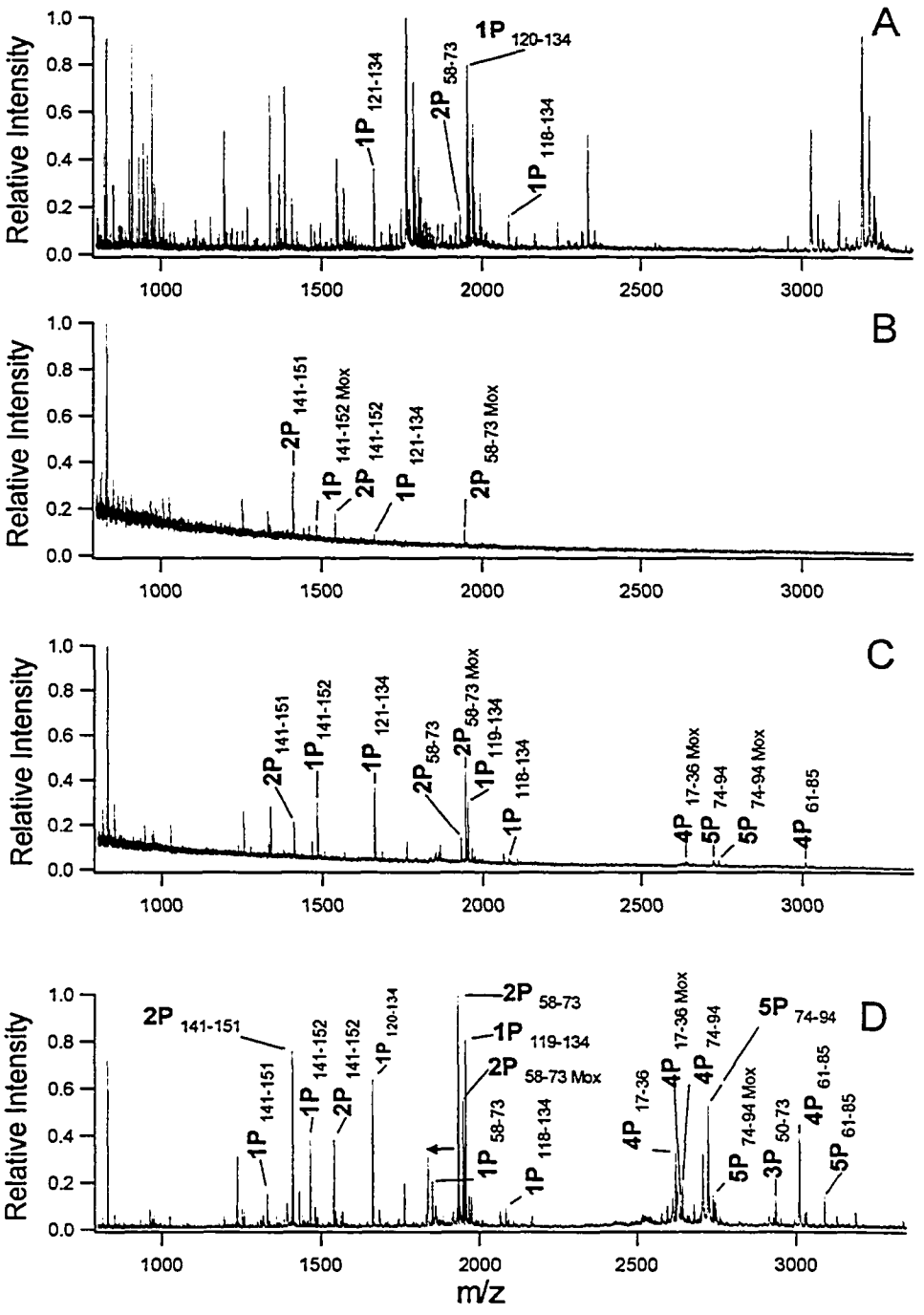

Figure 2.13. Detection sensitivity comparison between the commercially available packed pipettor tip IMAC from Millipore, OT-IMAC, and OTC-IMAC gauged by MALDI-MS with DHB matrix. A 1 pmol tryptic digest of α-casein was used for this purpose. A: digest without IMAC enrichment; B: Millipore-IMAC enriched digest; C: OT-IMAC enriched digest; D: OTC-IMAC enriched digest.

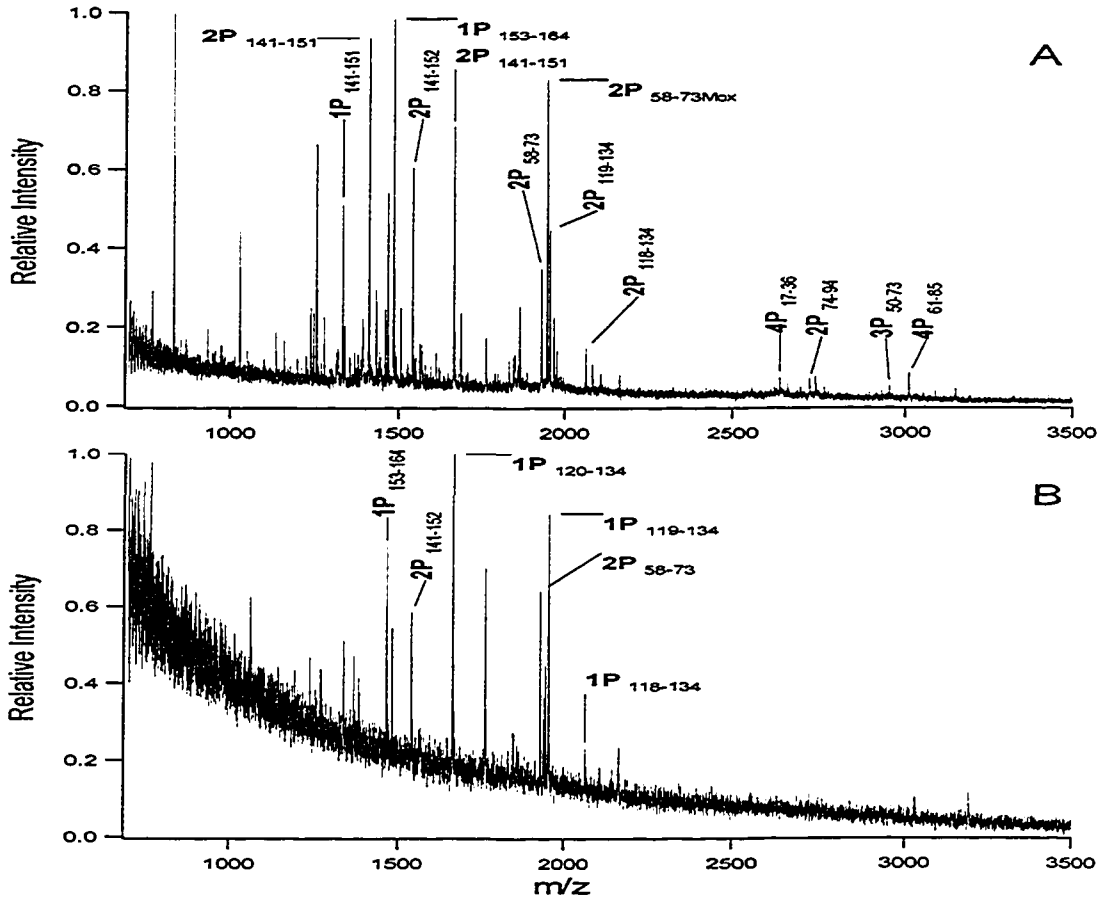

Figure 2.14. Analysis of 100 fmol (A) and 20 fmol (B) tryptically digested α-casein with OTC-IMAC.

2.3.7. The ABGRF-PRG03 Survey Sample

The ABGRF Proteomics Research Group conducted a study in 2003[35] to determine how many laboratories in the world were capable of successfully implementing methodologies for phosphoprotein analysis. For this purpose, a sample containing 5 pmol of an unphosphorylated protein digest containing 1 pmol of two synthetic phosphopeptides from the same protein, and in addition 200 fmol of a minor protein component were sent out to 54 laboratories in various locations. 3/54 laboratories correctly

identified both sites of phosphorylation. Of those using IMAC as an method for phosphopeptide enrichment, 1 out of 13 were successful at locating the phosphorylation site of one of the two phosphopeptides. Also of note is the fact that 8/13 labs using IMAC used the commercial product from MilliPore. Although our laboratory was not one of those participating in this test, we obtained the original sample in order to see how our IMAC would have fared.

With the application of OTC-IMAC, we were able to sequence 1 of the 2 phosphopeptides present in the mixture; this data is shown in Fig. 2.15. Without IMAC enrichment, we see a complex set of peaks in the spectrum, none of them corresponding to the phosphopeptides. After enrichment, we see that most of the background has been cleaned up, leaving the phosphopeptide of m/z 964.0 as the dominant peak in the spectrum. With MS/MS analysis of this peptide, we were then able to also assign the phosphorylation site as lying on Serine-3. We were however unable to detect the second phosphopeptide present, although the publication of the ABRF-PRG03 study shed some light on the possible reasons why. The sequence of the first peptide was SVpSDYEGK, whereas that of the second THILLFLPKpSVSDYEGK. The second peptide contains a missed trypsin cleavage site C-terminal to the K at position 9. Therefore, as rationalized by the authors of this study, thawing the sample may have resulting in tryptic cleavage of the shorter

peptide to pSVSDYEGK, a peptide with the same mass as the first phosphopeptide, albeit with a different phosphorylation site. This would in turn compromise the yield of this peptide, although after reanalyzing the MS/MS data we did not find evidence to support the presence of two phosphorylation sites. In conclusion, our IMAC is at least as good as the newest most successful product on the market (The only success with IMAC was using the Pierce Phosphopeptide Isolation Kit), and far better than one of the most popular products on the market.

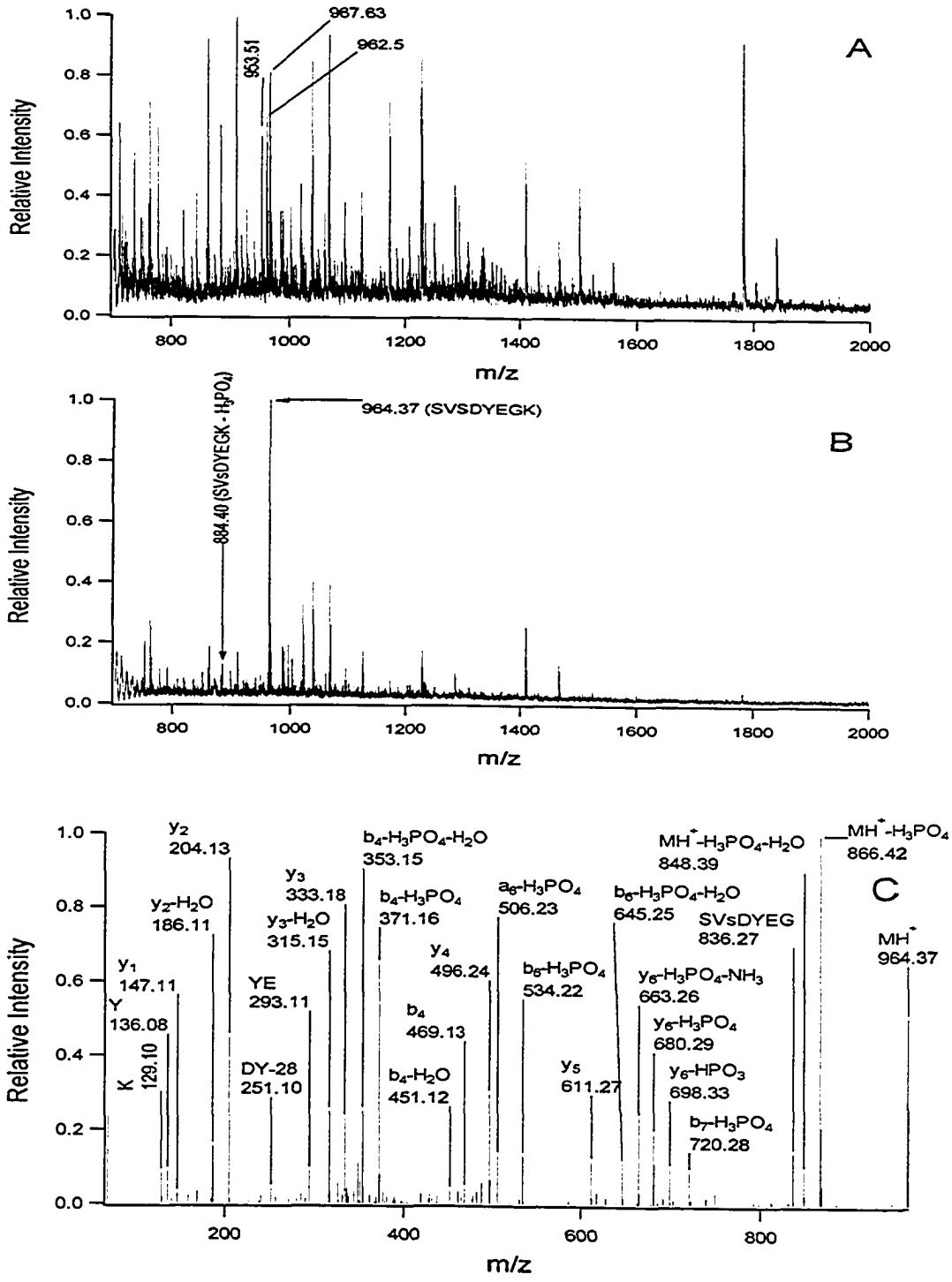

Figure 2.15. MALDI-MS analysis of the ABRF-PRG03 sample. A: MALDI-MS spectrum obtained prior to OTC-IMAC enrichment; B: MALDI-MS spectrum obtained after OTC-IMAC enrichment; C: MALDI-MS/MS spectrum of the phosphopeptide of m/z 964.4 (SVSDYEGK) shown in B.

2.3.8. Derivatization Chemistry of Carboxylate Groups to Eliminate Non-Specific Binding to OTC-IMAC.

Ficarro *et al.*[12] developed a methylation procedure for derivatization of carboxylate groups in order to eliminate non-specific binding to IMAC, which is believed to occur through the carboxylate group. We tested this procedure in an effort to improve the selectivity of the IMAC protocol. Through our initial experiments we saw that in addition to the methyl groups attached to the carboxylate groups of the phosphopeptides, another side reaction was taking place. Methylation of carboxylic acids results in a m/z increase of 14 Da for each carboxylate group which has been esterified, meanwhile mass shifts of +15 Da were being observed in our spectra. We hypothesized that besides glutamic and aspartic acid (these are the only amino acids which putatively react with the methanolic HCl used for the derivatization procedure) perhaps glutamine and asparagine were being esterified as well. The deamidation reaction and shown in Fig.2.16 has been hypothesized as a possible side reaction from the harsh conditions used for esterification of proteins and peptides, although to our knowledge no report exists to shown that it was indeed taking place. Evidence to support this explanation is shown in Fig.2.17.

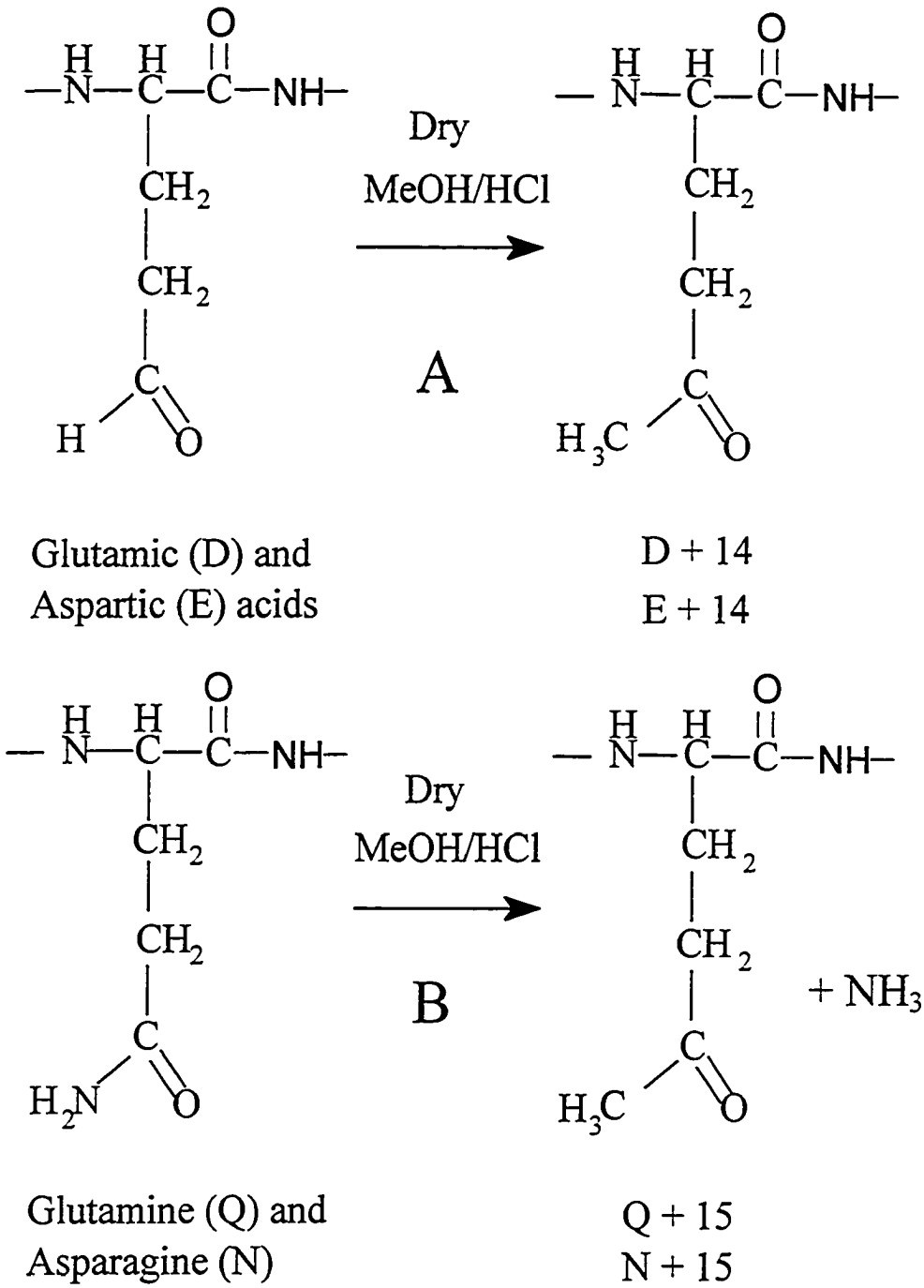

Figure 2.16. The chemistry of the methylation procedure used for esterifying carboxylic acids in proteins and peptides. A: The intended reaction of methanolic HCl with glutamic and aspartic acids of proteins and peptides; B: The deamidation side reaction of glutamine and asparagine with methanolic HCl.

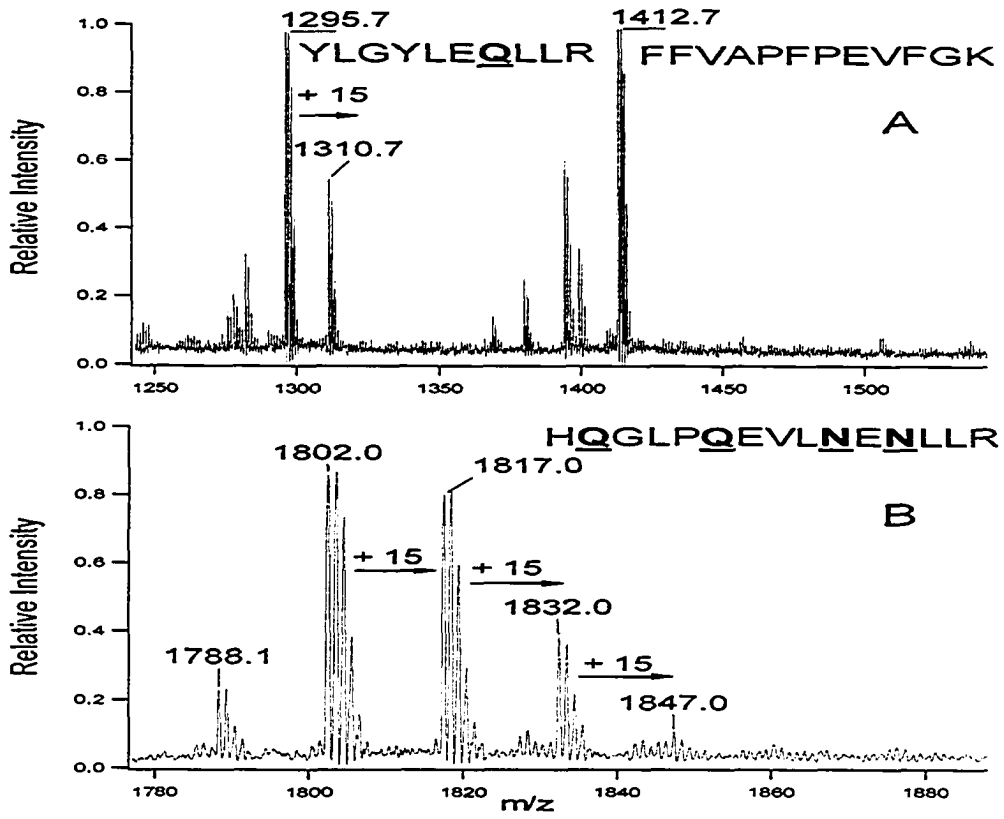

Figure 2.17. Evidence to support the deamidation reaction and subsequent esterification of glutamine (Q) and aparagine (N) residues. Expanded molecular ion region of the derivatized tryptically digested α-casein: A: peptide YLGYLEQQLLR at m/z 1295.7 and B: peptide HQGLPQEVLNENLLR at m/z 1802.0.

In the above figure, the peptide YLGYLEQLLR at m/z 1267.7, bearing one glutamine (Q) residue, is observed as the peak at m/z 1295.7 following methylation of the one glutamic acid and the C-terminus; and the peak at m/z 1310.7 showing a 15 Da increment corresponding to a modified asparagines residue. Note the absence of a plus-15 Da product in the FFVAPFPEVFGK peptide at m/z 1412.7. The bottom panel of the figure shows another peptide, HQGLPQEVLNENLLR (m/z 1759.9) when underivatized, shown here at m/z 1802.0 following methylation of the C-

terminus and two aspartic acid residues. This peptide contains two glutamine and two asparagines residues, hence we would expect up to a total of four plus-15 Da products. The fact that a distribution between three of them is observed, whereas peaks for not yet fully derivatized glutamic and aspartic residues are only very minor, indicates that the side reaction may be kinetically unfavourable, which is further supported by other work in our lab involving this particular derivatization: letting the reaction proceed for up to 24 hours leads to a less disperse distribution; however, we have not observed complete derivatization of all asparagines and glutamine residues. In any case, recognition of these minor modifications is important for peak assignment in the MALDI-MS spectra of derivatized digests.

In our initial experiments with the derivatization chemistry, we were only successful at derivatizing bulk quantities of α-casein. Ficarro *et. al.* used 100 uL of methanolic HCl for derivatization of 500 ug of protein from whole cell extracts of yeast. It was then observed that the volumes of reagents had to be scaled down in order to successfully derivatize small quantities of α-casein and detect the resulting peptides with good sensitivity. Figure 2.18 shows the trend we observed. The reason for this trend is not exactly clear, since the highest purity solvents available were used for the methylation reaction; however, intuitively it makes sense, since when dealing with small quantities of peptides a large volume of reagent will inevitably lead to more losses of peptides on walls of containers during

vacuum-drying, and introduce larger amounts of trace contaminants into our reaction vessel.

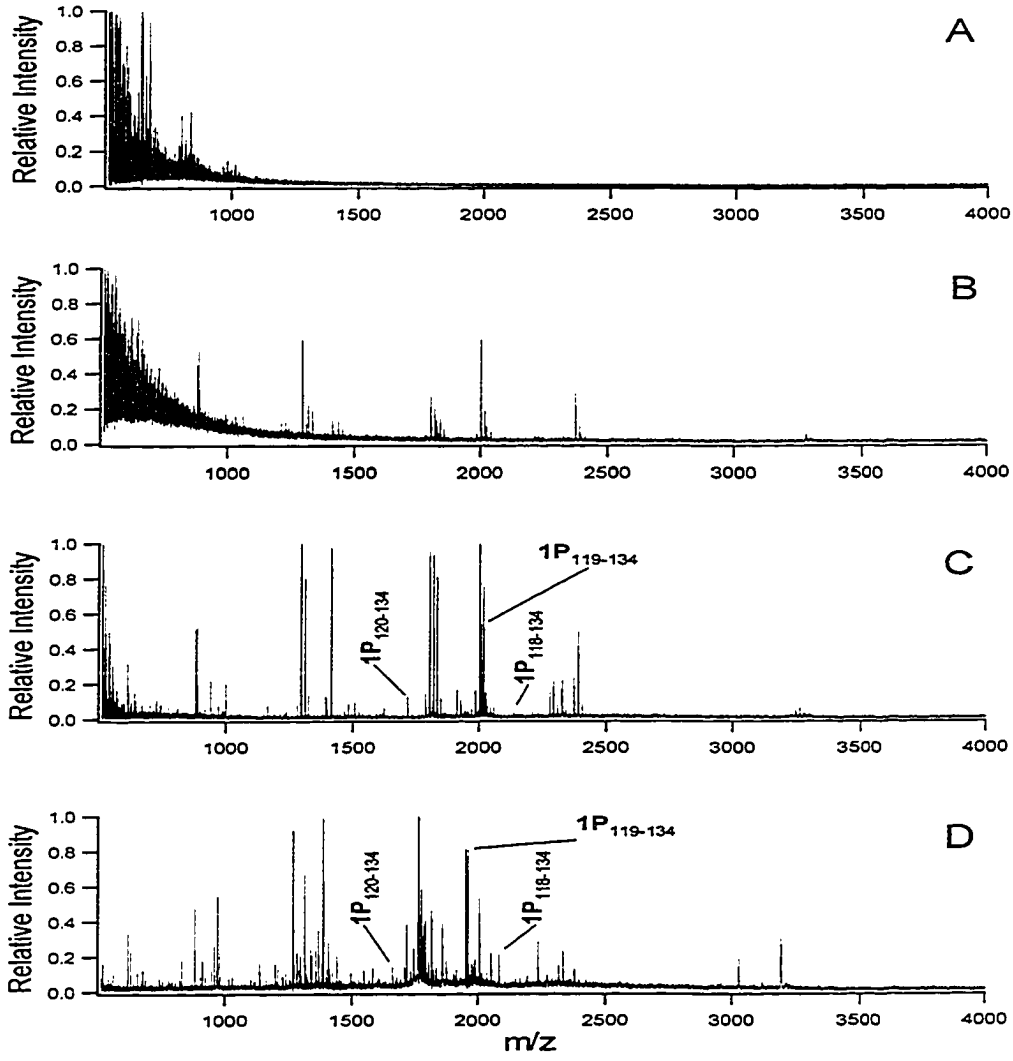

Figure 2.18. Effect of changing the derivatization reagent volume on derivatization efficiency for the methylation of carboxylate groups of tryptic peptides from 50 ng (2 pmol) of α-casein. Derivatization performed in 2 M methanolic HCl using: A: 100 μL; B: 10 μL; C: 1 μL of the reagent. Panel D shows the spectrum obtained by analyzing 2 pmol of an α-casein digest prior to derivatization.

2.3.9. Negative Ionization MALDI-MS Combined with Derivatization Chemistry.

During the course of working out the derivatization chemistry, it occurred to us that once a phosphopeptide has had its carboxylate groups esterified, the only acidic group left on the entire peptide would be the phosphate. Because phosphate is now the only group able to easily give up a proton, this should in turn make it the only ionizable group in negative ion mode MALDI-MS. This scheme is shown in Fig. 2.19.

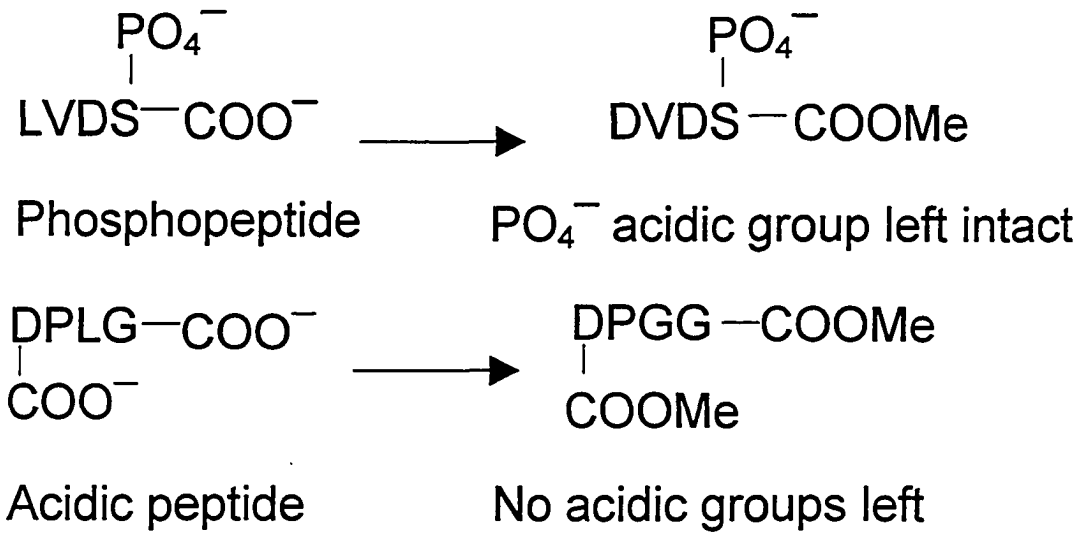

Fig. 2.19. The expected charge state of peptides following methyl esterification.

We were initially hoping for such a mode of analysis to also allow us to overcome ion suppression and replace IMAC altogether by a very inexpensive chemical procedure. However, this proven not to be the case: derivatization chemistry allows us to ionize phosphopeptides in a nearly-selective manner by applying the negative ion mode, however IMAC enrichment is still needed to detect low abundance phosphopeptides. Fig.

2.20 shows the results obtained from in-solution tryptic digests of α-casein following methyl esterification of carboxylate groups, with and without OTC-IMAC enrichment. As we see in comparing panels A and B, application of the negative ion mode to the esterified digest indeed allows us to ionize phosphopeptides in a nearly-selective manner; one extra phosphopeptide not observed in the positive ion mode even becomes detectable. The advantage of this approach lies in the fact that it allows us to clean up the spectrum, removing the great majority of non-phosphorylated peptides. With further enrichment by OTC-IMAC (panel C and D), we observed a greater total number of phosphopeptides, owing to the fact that we've captured and preconcentrated them into a very small area for detection by MALDI-MS. We found that some non-phosphorylated peptides still bind to IMAC even after esterification and extensive washing—presumably through hydrophobic interactions; application of the negative ion mode however, can then be used to further eliminate them from the spectrum, as shown in panel D.

In comparing the esterified and non-esterified digests enriched by OTC-IMAC, we see that with the esterification procedure some sensitivity is lost, as a lesser number of phosphopeptides was observed than that shown in Fig. 2.13D for the un-esterified α-casein digest. The intensities for the multiply phosphorylated peptides are also lower. We also found that in order for the derivatization to go to go to completion, samples must

first be desalted by way of a C_{18} Zip-tip, digests being typically carried out in a pH 8 sodium bicarbonate buffer, which could possibly account for at least some of the losses.

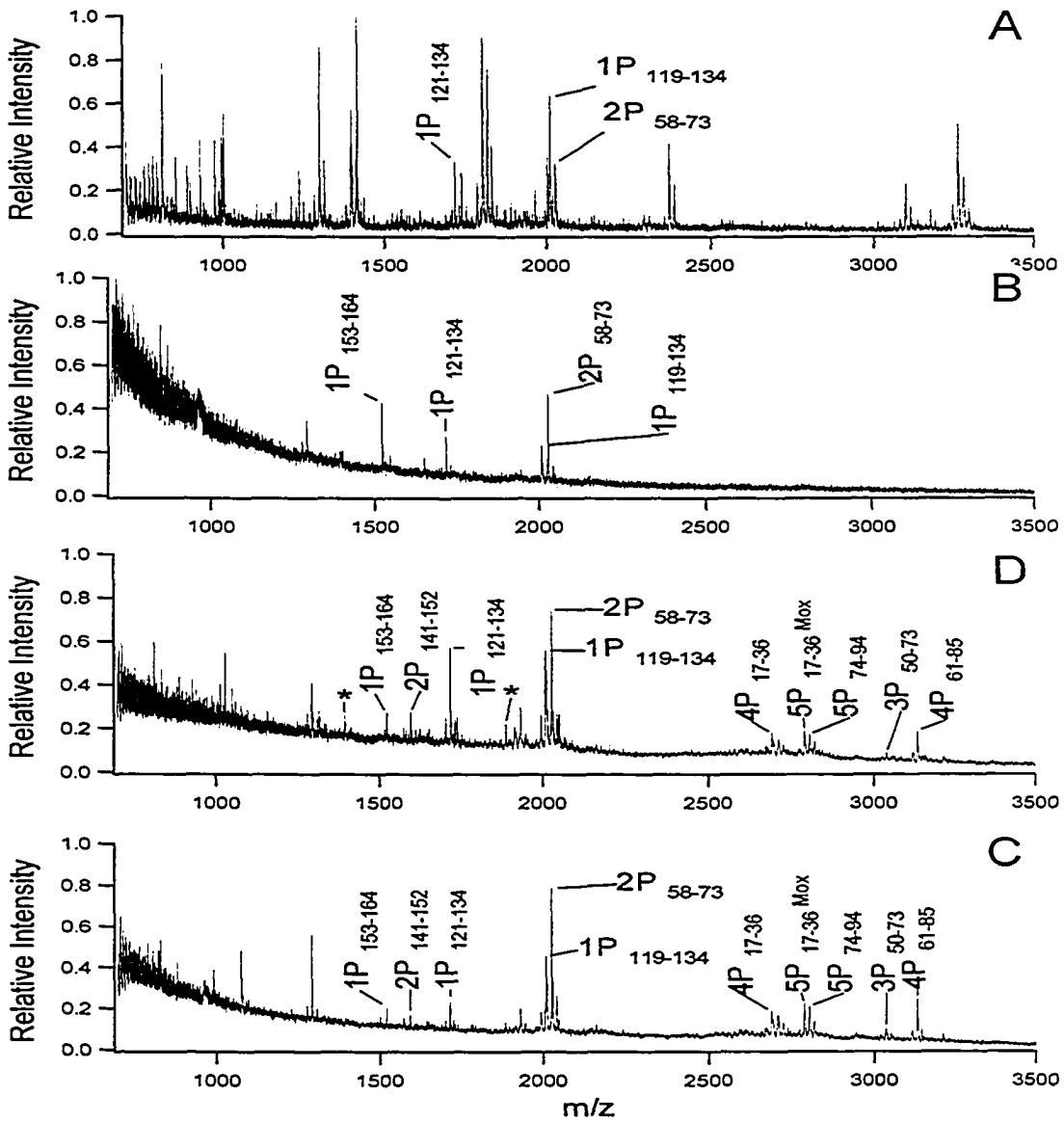

Figure 2.20. Effect of methyl esterification on ionization of phosphopeptides and OTC-IMAC enrichment. All spectra represent a 1 pmol solution digest of α-casein with the carboxylate groups methyl esterified. A: Unenriched digest, positive-ion mode; B: Unenriched digest, negative-ion mode; C: OTC-IMAC-enriched digest, positive-ion mode; D: OTC-IMAC-enriched digest, negative-ion mode. Asterisks: non-phosphorylated α-casein peptides.

2.3.10. Gel Experiments.

Since gel electrophoresis is currently one of the most widely used separation methods for displaying the proteome of a group of cells, we wanted to gauge the usefulness of our methods for proteins separated in this manner. Using the approach described in the preceding section, we were able to detect α-casein phosphopeptides in the negative mode down to a level of 4 pmol (100 ng) protein loaded on the gel. This data is shown in Fig. 2.21.

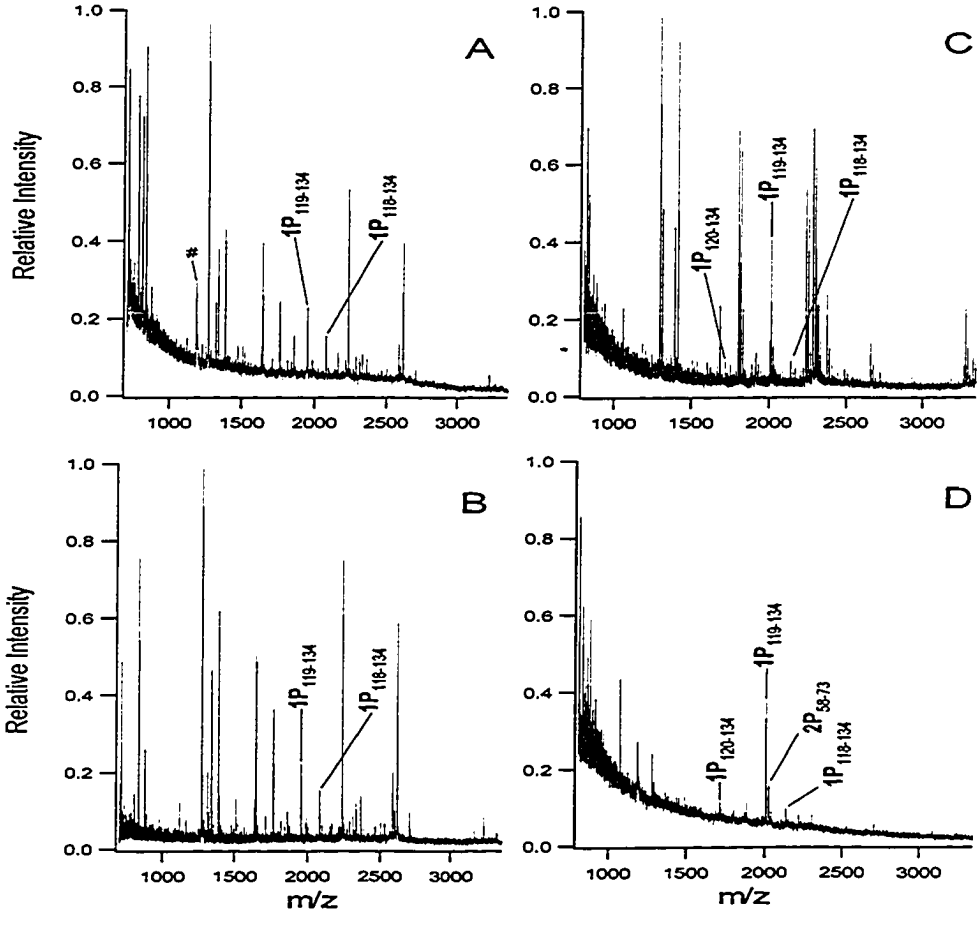

Figure 2.21. Maldi-MS spectrum of methyl-esterified α-casein peptides obtained from 4 pmol protein separated by SDS-PAGE. A: Native digest control, positive-ion Maldi-MS; B: Native digest control, negative-ion Maldi-MS; C: Derivatized digest, positive-ion Maldi-MS; D: Derivatized digest, negative-ion Maldi-MS.

It should be borne in mind that MS analysis of peptides obtained proteins digested enzymatically in-gel is typically much less sensitive than that of the same peptides obtained from an in-solution digest. Besides the lowered digestion efficiency in-gel, a number of extra steps are required for this kind of analysis. Peptides must be extracted several times from the gel, they must be then be vacuum-dried to concentrate them, which leads to losses of peptides on walls of containers as they tend to be rather sticky; and finally they must also be desalted since the ammonium bicarbonate buffer used for the digestion also becomes concentrated during vacuum drying. Hence, gauging the merits of a method is typically done through analyzing gel-separated proteins.

Typically, 10-30 ng is the low-end limit of detection for Coomassie-blue stained protein bands in SDS-PAGE. This is also the low-end limit for MS analysis of in-gel digested proteins. In order to show that we are able to enrich phosphopeptides at these low levels, we ran SDS-PAGE gels with α-casein; this was followed by extraction of peptides, concentration, as well as finally a desalting step on miniaturized C_{18} columns (Zip-Tip). Fig. 2.22 shows the results obtained for 10 ng (~400 fmol) α-casein enriched by OTC-IMAC with, and without the prior derivatization of carboxylate groups. As is evident in the figure, phosphopeptide enrichment with OTC-IMAC is still feasible even at these low levels. As discussed in the preceding section for in-solution digested α-casein, application of the negative ion mode once again seems to favour ionization of, at least in this case, one additional phosphopeptide which was otherwise undetectable. In comparing panels E and F, we see that the negative

ion mode allows us to eliminate some unphosphorylated trypsin autolysis peaks and also some non-specifically binding α-casein peptides. The presence of these peaks was somewhat puzzling at first, since non-specific binding to IMAC is generally believed to take place through the carboxylate group; however, non-specific binding of fully esterified peptides to IMAC has also been observed by others[13], who found that the use of a NaCl wash had no effect on efforts to wash away esterified peptides from commercially available IMAC beads. Presumably, hydrophobic interactions are responsible for this type of non-specific binding, and eliminating it completely has proved to be rather inconsistent, owing perhaps to the nature of chemical surface phenomena, which in general are quite difficult to predict.

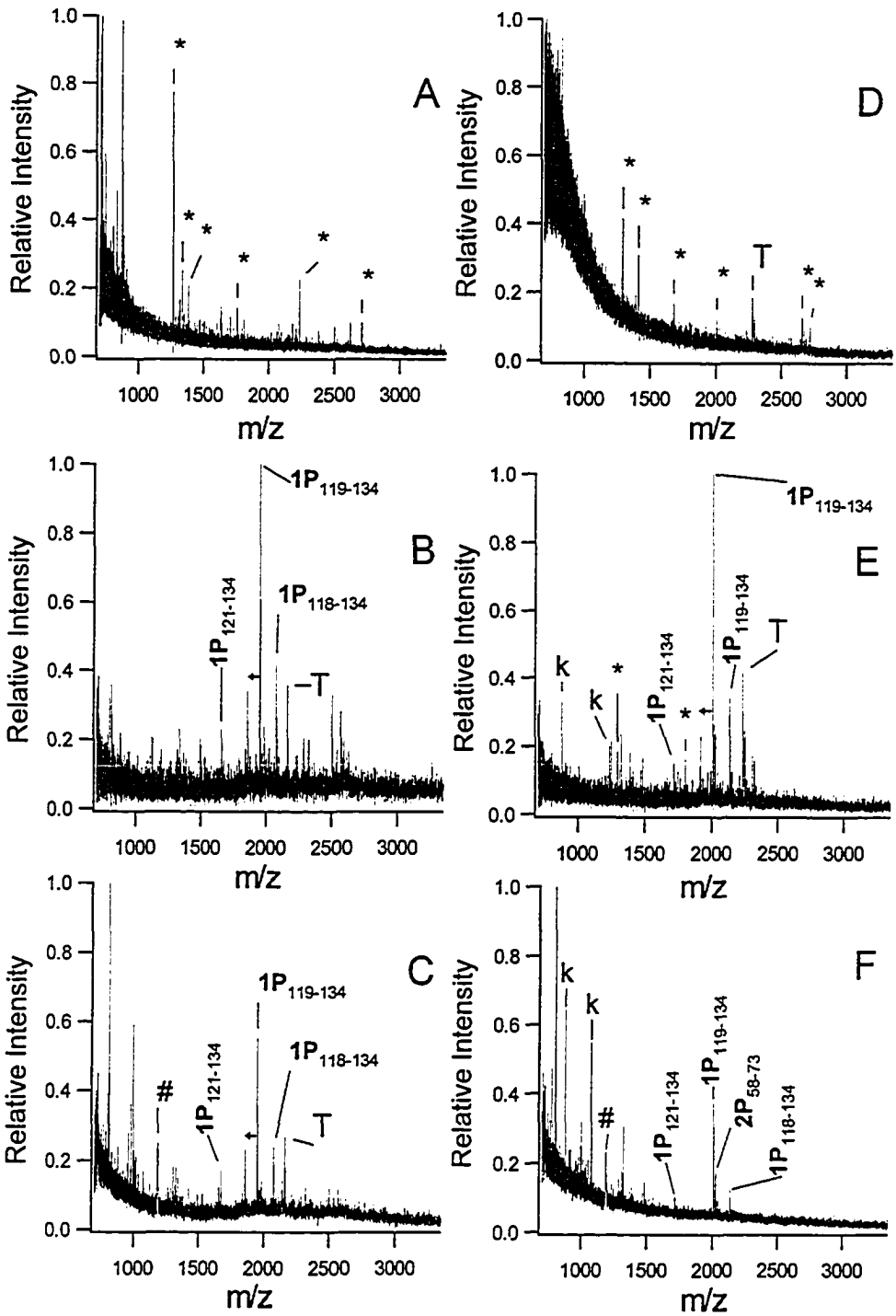

Fig. 2.22. 10 ng α-casein separated by SDS-PAGE, digested with trypsin, and enriched by OTC-IMAC. A: Underivatized digest without IMAC enrichment; B: Underivatized digest enriched with OTC-IMAC, positive ion mode MALDI-MS. C: Underivatized digest enriched with OTC-IMAC, negative ion mode; D: Derivatized digest without IMAC enrichment; E: Derivatized digest enriched with OTC-IMAC, positive ion mode; F: Derivatized digest enriched with OTC-IMAC, negative ion mode.

2.3.11. Conclusions

Open tubular IMAC and open tubular capillary IMAC are more effective and sensitive at enriching phosphopeptides than one of the most popular IMAC products from Millipore. Sensitivity for IMAC enrichment and detection of phosphopeptides is highest using capillary IMAC, with DHB MALDI-MS matrix. The esterification chemistry developed by Ficarro *et. al.*[12] can be scaled down to handle sub-pmol amounts of proteins separated by SDS-PAGE. We've also shown that application of negative-ion MALDI-MS for esterified peptides leads to great improvements in ionization selectivity, and allows elimination of most non-phosphorylated peptides from MS-spectra.

2.3.12. References Cited

1. Rudolph, J., Tolliday, N., Schmitt, C., Schuster, S.C., Oesterhelt, D., **1995**, *EMBO J.*, 14(17), 42494259.
2. Cantrell, D., **1996**, *Annu. Rev. Immunol.* , 14, 259-
3. Cohen, P. 1992, Trends Biochem. Sci., 17, 408.
4. Hunter, T., 2000, Cell, 100, 113-
5. MacKintosh, C. **1998**, *Curr. Opin. Plant Biol.*, 1, 224-229.
6. Darnell, J.E.Jr., **1997**, *Science*, 277, 1630-1635.
7. Pawson, T., Scott, J.D., **1997**, *Science*, 278, 2075-2080.
8. Ma, Y., Lu, Y., Zeng, H., Rog, D., Mo, W., Neubert, T.A., 2001, Rapid Commun. Mass Spectrom., 15, 1693-1700.

9. Bonewald, L.F., Bibbs, L., Kates, S.A., Khatri, A., McMurray, J.S., Medzihradszky, K.F., Weintraub, S.T., 1999, J. Peptide Res., 161-169.

10. Janek, K., Wenschuh, H., Bienert, M., Krause, E., **2001**, *Rapid Commun. Mass Spectrom.*, 15, 1593-1599.

11. Neubauber, G., Mann, M., 1999, Anal. Chme., 71, 235-242.

12. Ficarro, S.B., McCleland, M.L., Stukenberg, P.T., Burke, D.J., Ross, M.M., Shabanowitz, J., Hunt, D.F., White, F.M., **2002**, *Nature Biotech.*, 20, 301-305.

13. Haydon, C.E., Eyers, P.A., Aveline-Wolf, L.D., Resing, K.A., Maller, J.L., Ahn, N.G., **2003**, *Mol. Cell. Proteomics*, 2, 1055-1067.

14. Nuhse, T.S., Stensballe, A., Jensen, O.N., Peck, S.C., **2003**, *Mol. Cell. Proteomics*, 2, 1234-1243.

15. Whittal, R.M., Keller, B.O., Li, L., **1998**, *Anal. Chem.*, 70, 5344-5347.

16. Keller, B.O., Li, L., **2000**, *J. Am. Soc. Mass Spectrom.*, 11, 88-93.

17. Keller, B.O., Li, L., **2001**, *Anal. Chem.*, 73, 2929-2936.

18. Keller, B.O., Li, L., **2001**, *J. Am. Soc. Mass, Spectrom.*, 12, 1055- 1063.

19. Shevchenko, A., Jensen, O.N., Podtelejnikow, A.V., Sagliocco, F., Wilm, M., Vorm, O., Mortensen, P., Shevchenko, A., Boucherie, H., Mann, M. **1996**, *Proc. Natl. Acad. Sci. USA*, 93, 14440-14445.

20. Anspach, F.B., 1994, J. Chromatogr. A, 672, 35-49.

21. Gimpel, M., Unger, K., **1982**, *Chromatographia*, 16, 117-125.

22. Zhang, N., Doucette, A., Liang, Li, **2001**, *Anal. Chem.*, 73, 2968-2975.

23. Liu, H., Stupak, J., Zheng, J., Keller, B.O., Brix, B.J., Fliegel, L., Li, L., **2004**, *Anal. Chem.*, 76, 4223-4232.

24. Hart, S.R., Waterfield, M.D., Burlingame, A.L., Cramer, R., **2002**, *J. Am. Soc. Mass Spectrom.*, 13, 1042-1051.

25. Posewitz, M.C., Tempst, P., **1999**, *Anal. Chem.*, 71, 2883-2892.

26. Brill, L.M., Salomon, A.R., Ficarro, S.B., Mukherjii, M., Stettler-Gill, M., Peters, E.C., **2004**, *Anal. Chem.*, 76, 2763-2772.

27. Stensballe, A., Andersen, S., Jensen, O.N., **2001**, *Proteomics*, 1, 207-222.

28. Li, S., Dass, C., **1999**, *Anal. Biochem.*, 270, 9-14.

29. Ren, D., Penner, N.A., Slentz, B.E., Inerowicz, H., Rybalko, M., Regnier, F.E., **2004**, *J. Chrom. A*, 1031, 87-92.

30. Andersson, L., **1991**, *J. Chromatogr.*, 539, 327-334.

31. Shoemaker, M.T., Haley, B.E., **1993**, *Biochemistry*, 32, 1883-1890.

32. Andrecht, S., Schwammle, L., Ruter, C., Seler, A., Matheis, D., Hendriks, R., Anders, J., *Proceedings of the 52st ASMS Conference on Mass Spectrometry and Allied Topics*, Nashville, TN, June 2004; Poster 404.

33. Campbell, D.G., Morrice, N.A., **2002**, *J. Biomol. Tech.*, 13, 119- 130.

34. Voisin, S., Watson, D., Young, M., Ding, W., Tessier, L., Kelly, J., Andrecht, S., Schwammle, L., Ruter, C., Seler, A., Matheis, D., Hendriks, R., Anders, J., *Proceedings of the 51st ASMS Conference on*

Mass Spectrometry and Allied Topics, Montreal, QC, June 2004; Poster 247.

35. Arnott, D., Gawinowicz, M.A., Grant, R.A., Neubert, T.A., Packman, L.C., Speicher, K.D., Stone, K, Turck, C.W., **2003**, *J. Biomol. Tech.*, 14, 215.

Chapter 3

Application of OT-IMAC and OTC-IMAC to the *in Vitro* Phosphorylated Na^+/H^+ Exchanger Protein His182.[a]

3.1 Introduction

In this chapter, we will describe the application of our OT-IMAC and OTC-IMAC in combination with sequential enzyme digestion followed by MALDI-MS and MALDI-MS/MS, for enrichment and identification of phosphorylation sites from the Na^+/H^+ exchanger fusion protein, His182, which was phosphorylated *in vitro* with the kinase Erk2. His182 is a pH regulatory protein which removes excess intracellular acids and is involved in cell growth, differentiation, and cell migration[1]. Its protein sequence is shown in Fig. 3.1. Residues 2-183 correspond to amino acids 634-815 of the NHE1 isoform of the human Na^+/H^+ exchanger, an initiator methionine residue was necessary at the N-terminus, and a hexahistidine tag was fused at the C-terminus to allow for protein purification. The protein used in this study includes all the amino acids known to be involved in regulation of the protein by phosphorylation.[2,3] Deduction of the amino acids targeted by the kinase Erk2 is believed to be crucial for understanding the physiological mechanism of activation and regulation of the protein.

[a] A form of this chapter and parts of chapter 2 was submitted for publication: J. Stupak, H. Liu, Z. Wang, B. Brix, L. Fliegel, L., L.Li, "Nanoliter Sample Handling Combined with Chemical Derivatization and Negative Ion MALDI-MS Analysis for Detection of Gel-Separated Phosphoproteins". Dr. Liu prepared the OT-IMAC enrichment of His182 and collected the MS/MS spectra. L. Fliegel and B. Brix provided the His182 samples.

1	11	21	31
MILRNNLQKT	RQRLRSYNRH	TLVADPYEEA	WNQMLLRRQK
41	51	61	71
ARQLEQKINN	YLTVPAHKLD	SPTMSRARIG	SDPLAYEPKE
81	91	101	111
DLPVITIDPA	SPQSPESVDL	VNEELKGKVL	GLSRDPAKVA
121	131	141	151
EEDEDDDGGI	MMRSKETSSP	GTDDVFTPAP	SDSPSSQRIQ
161	171	181	
RCLSDPGPHP	EPGEGEPFFP	KGQHHHHHH	

Figure 3.1. Amino acid sequence of His182.

3.2. Experimental

3.2.1 Materials.

Bovine trypsin, Endoproteinase Asp-N (from *Pseudomonas fragi* mutant strain), Endoproteinase Glu-C (from *Staphylococcus Aureus*) were obtained from Sigma Aldrich Canada (Oakville, ON). 4-hydroxycinnamic acid (HCCA), 2,5-dihydroxybenzoic acid (DHB), acetyl chloride, 2-[N-Morpholino]ethanesulfonic acid, iron(III) chloride, ammonium dihydrogenphosphate, ammonium bicarbonate, hydrochloric acid, trifluoroacetic acid, glacial acetic acid, 3-glycidoxypropyltrimethoxysilane, iminodiacetic acid, sodium hydroxide, and all solvents used were of the highest available purity, and were from Sigma Aldrich Canada. HCCA was purified by recrystallization from ethanol prior to use. Deionized water was from NANOpure water system (Barnstead/Thermolyne).

3.2.2 Protein Digests and Electrophoresis.

In-gel tryptic digests were performed using the method of Shevchenko *et al.*[29] as outlined in section 2.2.2.

3.2.3 Expression and Purification of Fusion Proteins.

The carboxyl-terminal 182 acids of the human Na^+/H^+ exchanger (NHE1) was expressed as a fusion protein with a C-terminal hexahistidine tag (His182) using the plasmid pDest 14 and the Gateway Cloning System as described[4]. *E.coli* strain BL21-SI was induced with 0.3 M NaCl for 3 h. His182 protein was harvested using standard conditions, and the protein was purified from the supernatant of *E.coli* via Ni-NTA affinity chromatography according to manufacturer's directions (Qiagen). Residues 2-183 of this protein correspond to amino acids 634-815 of the NHE1 isoform of the human Na^+/H^+ exchanger[5]. An initiator methionine residue was necessary at the N-terminus, and a hexahistidine tag at the C-terminus was necessary for protein purification.

3.2.4 *In Vitro* Phosphorylation of His182.

The standard reaction conditions for phosphorylation of His-182 contained 10.0 ug of His182, 0.03 ug Erk2 (Biomol, Plymouth Meeting, PA), 12.5 mM 3-(4-morpholino)propane-sulfonic acid (MOPS) pH 7.2, 0.5 mM EGTA, 2 mM DTT, 8.5 mM magnesium chloride, 6 uM okadaic acid, 0.24 mM sodium fluoride, 500 uM ATP, and 1 uL of 10 uM ATP in a final volume of 24 uL^2. Samples were incubated at 30 °C for 90 min., and the

reaction was terminated by the addition of SDS-PAGE loading buffer. Samples were run on standard 0.5 mm 12% SDS-PAGE gels (BioRad).

3.2.5 Sequential Digestion of Phosphopeptides with Endoproteinase Asp-N.

Whenever phosphorylation sites could not be assigned from the MS/MS spectra, 10 uL of the ammonium phosphate eluate containing OT-IMAC enriched phosphopeptides from tryptically digested His182 was adjusted to pH 7.0 with 2% ammonia solution. One microliter of 20 ng/uL endoproteinase Asp-N in 100mM phosphate buffer pH 7.0 was then added. Digestion was allowed to proceed for 48 hours at 37 °C, and the products analyzed by MALDI-MS and MALDI-MS/MS.

3.2.6 Open Tubular IMAC Construction and Phosphopeptide Purification.

IMAC devices, both OT-IMAC and OTC-IMAC were constructed and used as described in section 2.2.3.

3.2.7 Mass Spectrometry.

MALDI was performed on a Voyager Elite reflectron time-of-flight mass spectrometer (Framingham, MA) equipped with a nitrogen laser (λ = 337 nm) by collecting signals averaged from 200-400 laser shots. Spectra were calibrated externally using monoisotopic ions of a pepcal standard, or internally using trypsin autolysis peaks. MALDI spectra were analyzed using the Igor Pro Software package (WaveMetrics, Lake Oswego, OR).

MALDI MS/MS spectra were collected on an Applied Biosystems MDS-Sciex QSTAR Pulsar QqTOF instrument equipped with an orthogonal MALDI source employing a 337 nm nitrogen laser (Missisauga, ON). The instrument was operated in positive-ion mode, and collision induced dissociation (CID) was carried out with argon as the collision gas. Peptide mass mapping and sequencing were done using UCSF ProteinProspector tools MS-Digest and MS-Product (http://prospectro.ucsf.edu/) for protein identification and peptide sequencing.

3.3. Results and Discussion

3.3.1 OT-IMAC Enrichment and Sequencing of Phosphorylation Sites in His182.

Prior to the development of OT-IMAC and OTC-IMAC, several unsuccessful attempts at sequencing the phosphorylation sites of the His182 protein were made by the Alberta Cancer Board proteomics resource laboratory staff. The commercial IMAC device from Millipore was used for this purpose. This was the impetus for development of the herein presented home-made IMAC devices. While the OTC-IMAC device discussed in Chapter 2 allows us to detect phosphopeptides at very low levels, actual sequencing of phosphorylation sites requires multiple experiments; in addition, the quadrupole-TOF instrument used for MS/MS requires longer acquisition times and more sample than a simple TOF instrument used for peptide detection and peptide mass mapping. As

suggested by Mann[6], one typically needs around 1 ug of tryptically digested protein from an SDS-PAGE gel for successful analysis of post-translational modifications. We obtained ~10 μg of phosphorylated His182, this being the total amount of protein loaded on an SDS-PAGE gel. Although the stoichiometry of phosphorylation is unknown, the reaction was carried out *in vitro*, and it would not be surprising if the stoichiometry was unrealistically high compared to the actual phosphorylation *in vivo;* still this is irrelevant for this analysis. OT-IMAC enrichment of phosphopeptides was performed in order to allow for sequencing of phosphorylation sites. Figure 3.2 shows the MALDI spectra of His182 digests with and without OTC-IMAC enrichment. As is evident in the figure, without OT-IMAC enrichment, phosphopeptide peaks are largely masked by a background of non-phosphorylated peptides. Following IMAC enrichment, the intensity of these peaks increases, and new phosphopeptides also become detectable owing to the overcoming of ion-suppression effects. As can be seen in the figure, for this particular protein non-specific binding is not a big concern, as the majority of peaks observed in the MALDI-MS spectra of IMAC-enriched digests are indeed phosphopeptides.

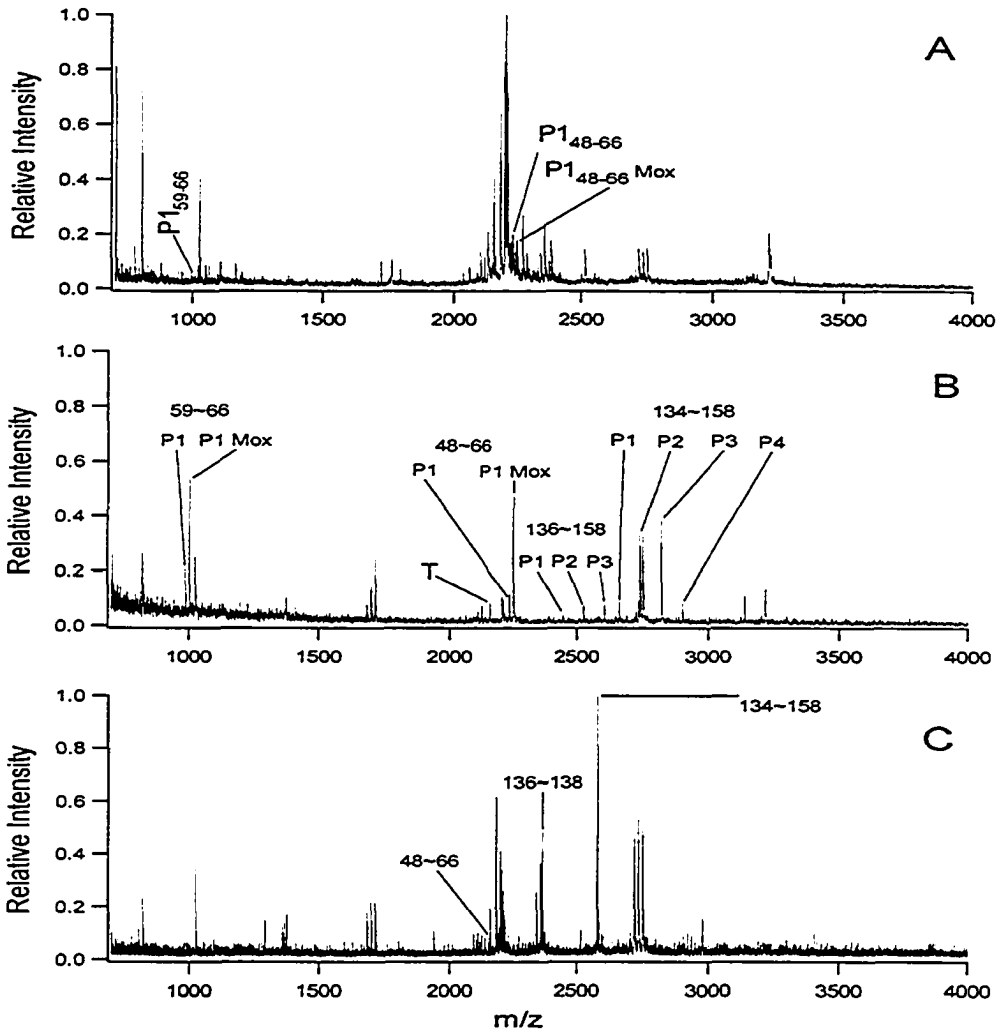

Figure 3.2. MALDI-MS analysis of His182. A: In-gel tryptic digest of His-182 before OT-IMAC treatment; B: MALDI-MS spectrum of OT-IMAC enriched His182; C: MALDI-MS spectrum of the OT-IMAC enriched control: His182 untreated with the kinase Erk2. M/z in Da: $P1_{59\text{-}66\ Mox}$ = 1002.4; $P1_{48\text{-}66}$ = 2237.1; $P1_{136\text{-}158}$ = 2445.1; $P2_{136\text{-}158}$ = 2525.1; $P1_{136\text{-}158}$ = 2605.1; $P1_{134\text{-}158}$ = 2660.1; $P2_{134\text{-}158}$ = 2740.1; $P3_{134\text{-}158}$ = 2820.1; $P1_{134\text{-}158}$ = 2900.1.

Shown in Figure 3.3A is the MS/MS spectrum of the peak at m/z 1002.4 from Figure 3.2. From the product ions generated through CID fragmentation, the phosphorylation site was mapped as lying on Ser-61 in the sequence 59-LDSPTMSR-66. The MS/MS spectrum shows a protonated molecular ion (MH^+), a loss of phosphoric acid, followed by a

series consisting mostly of y-type ions. From the presence of the y_6 ion and corresponding y_6-H_3PO_4 ion, the phosphorylation site can be said to lie either on Ser-61, or Thr-63. Since a loss of phosphoric acid from the y_5 ion was not observed, we can conclude with good certainty that it is Ser-61 which contains the phosphate group on this peptide, and not Thr-63. Also notable is the absence of the y_4 ion in this spectrum, not unexpectedly so since bond cleavage on the C-terminal side of Pro residues often shows reduced intensity, and is often absent.

In Figure 3.3B we have the MS/MS spectrum of the 2237.1 ion from Figure 3.2. The sequence of this peptide is 48-INNYLTVPAKSDSPTMSR-66, and it contains a tryptic missed cleavage site at K_{57}. There are five possible phosphorylation sites on this peptide. As in the foregoing discussion for the m/z 1002.4 ion, the phosphorylation site was unambiguously assigned as lying on Ser-61 from the presence of the y_6 ion, and because no loss of phosphoric acid was observed for the y_5 and y_4 ions. The presence of the b_{13} ion and lack of the corresponding b_{13} ion minus phosphoric acid, also confirms that there is no phosphorylation on residues 48-60. Therefore, fragmentation of the m/z 2237.1 ion further confirms our assignment of phosphorylation on Ser-61 obtained previously from the m/z1002 ion.

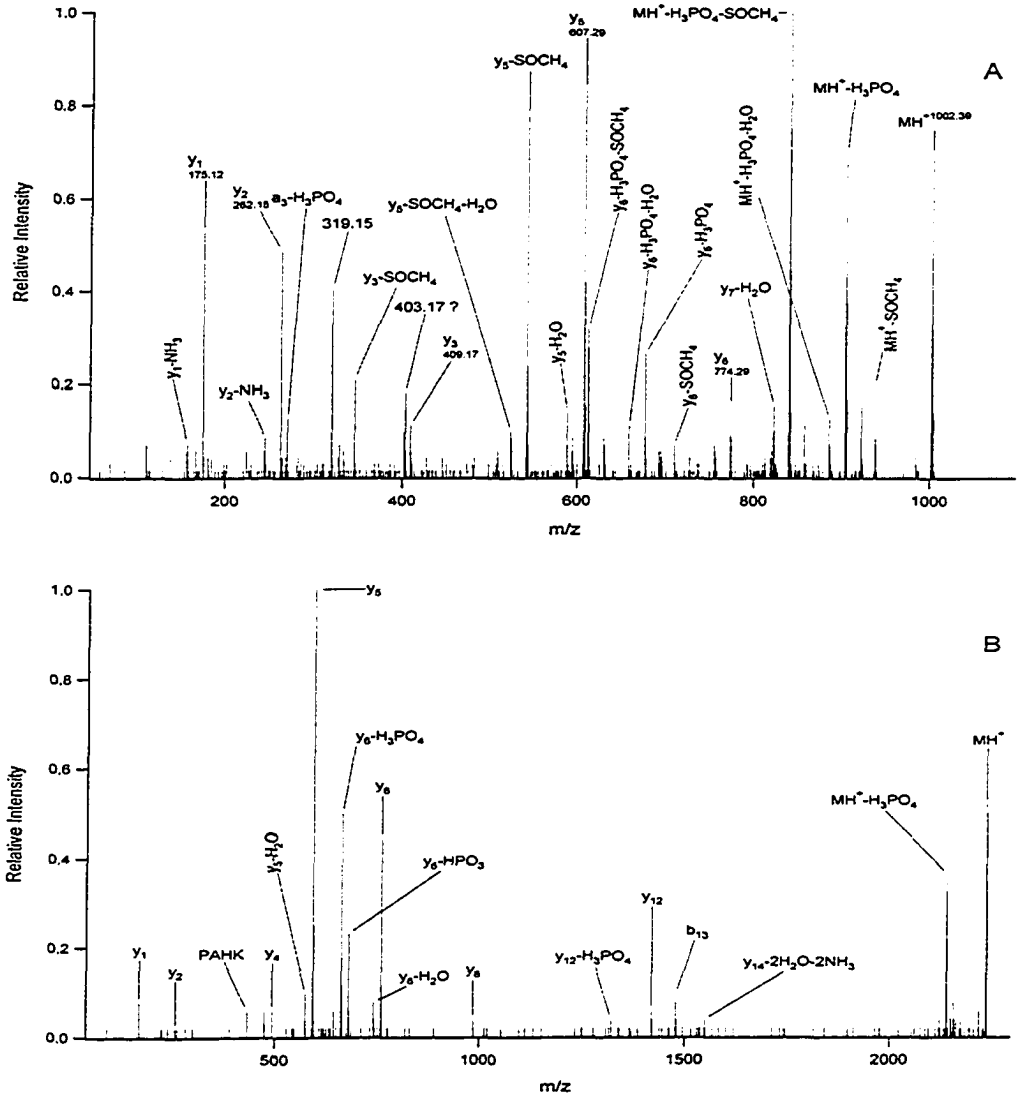

Figure 3.3. MALDI-MS/MS analysis of phosphopeptides of m/z 1002.4 and 2237.1 from figure 3.2. A: peptide 59-LDSPTMSR-66, m/z 1002.4. B: Peptide 48-INNYLTVPAHKLDSPTMSR-66, m/z 2237.1.

Figure 3.4A shows the MALDI-MS/MS spectrum obtained from the ion of m/z 2820.1 from figure 3.2. This peak corresponds to the peptide with the sequence 134-SKETSSPGTDDVFTPAPSDSPSSQR-158. We can see the molecular ion and the corresponding losses of 3 phosphoric acid residues, confirming that the peptide is triply phosphorylated. Since for the y_1 to y_5 series of ions no losses of phosphoric acid were observed, we conclude that the sequence PSSQR contains no phosphate residues. Since

the y_6 ion shows the characteristic loss of phosphoric acid, we therefore assign Ser-153 as phosphoserine. Since the y_{11} ion shows a loss of phosphoric acid, and because Ser-153 was determined to be phosphoserine, the possibility of Ser-151 being phosphorylated is now excluded. At this point we know that Ser-153 is phosphoserine and that Ser-151 is not; so the assignment of the y_{15} ion (144-DVFTPAPSDSPSSQR-158 containing 2 phosphates) with a loss of phosphoric acid indicates that the other phosphorylated residue in this sequence must be Thr-147. Next, since we assign the y_{19} ion (it's mass corresponds to 2 phosphates), and because the 2 phosphorylated residues were determined to be Ser-153 and Thr-147, we can also infer that Thr-142 is not phosphorylated. Since neither the y_{20} – y_{24} nor the b_1 – b_6 series ions were observed, we can only narrow down the last phosphate residue as lying on either a Ser or Thr somewhere in the sequence SKETSS. To resolve this ambiguity, endoproteinase Asp-N was used to further cleave the 2820.1 peptide into smaller fragments amenable to further CID fragmentation. Asp-N cleaves peptide bonds on the N-terminal side of glutamic acid (D), aspartic acid (E), and Cys (C) residues. After digesting the OT-IMAC eluate with Asp-N for 15 hours at 37 °C, the peak at m/z 1088.4 corresponding to the singly phosphorylated peptide 134-SKETSSPGT-142 was observed. Fragmentation of this peptide with the Q-TOF instrument gave the MS/MS spectrum shown in figure 3.4B.

In looking at Figure 3.4B, loss of a single phosphoric acid residue confirms that this peptide is singly phosphorylated. The presence of the b_2 and b_3 ions exclude the possibility of Ser-134 being phosphorylated; the y_4 ion, corresponding to the SPGT fragment without any phosphate resides, excludes the possibility of phosphorylation at Thr-142 and Ser-139. If we assign the

phosphate group as lying on Thr-137, the peak labeled KET in Figure 3.4B could not be assigned. From the presence of the b_5-H_3PO_4 peak, we can then conclude that it must be Ser-138 which is phosphorylated.

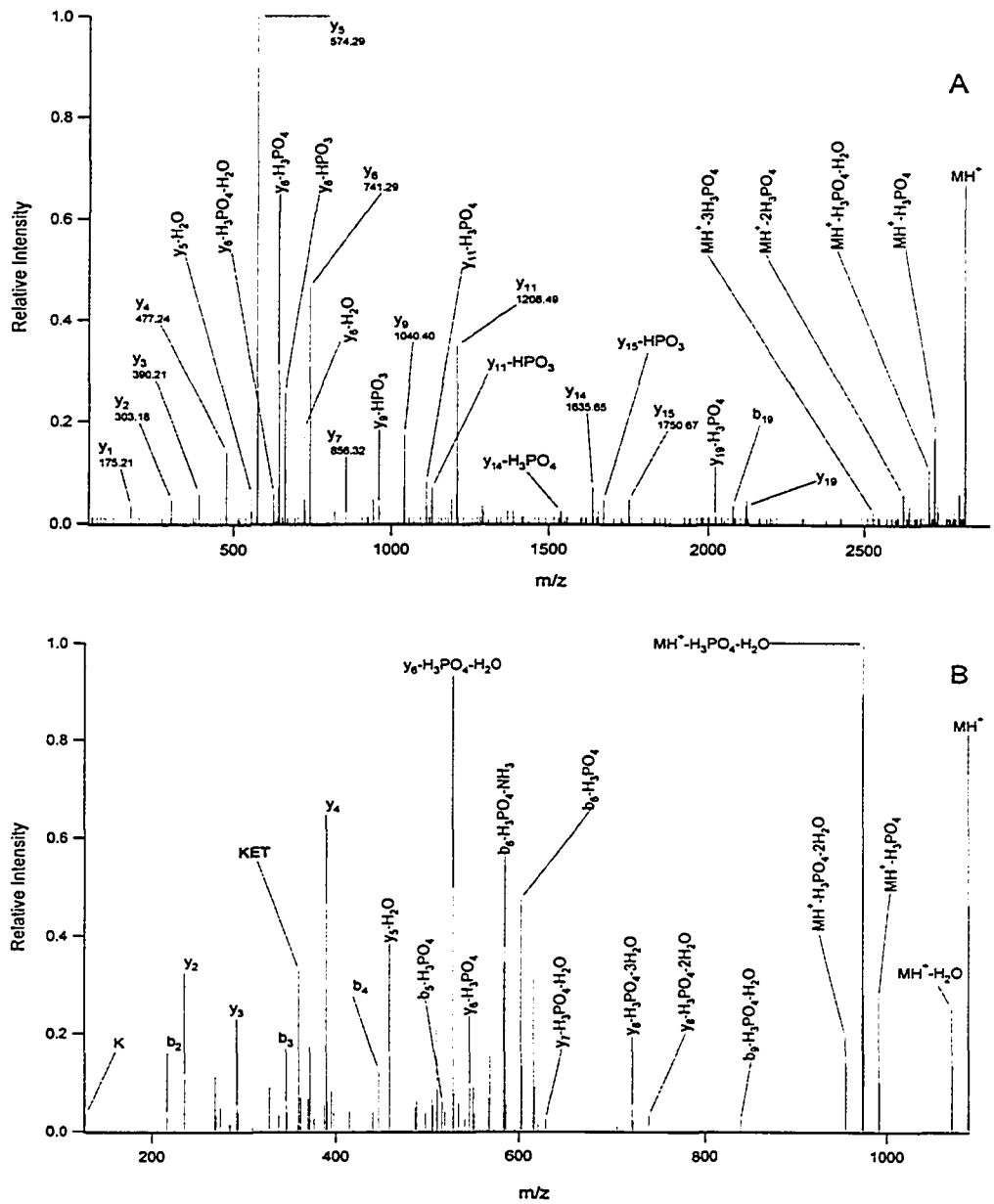

Figure 3.4. MALDI-MS/MS spectra of the m/z 2820.1 peak from Fig.3.2 and the smaller peptide resulting from sub-digestion of the said peak with the endopeptidase Asp-N. A: MALDI-MS/MS spectrum of the m/z 2820.1 peak from Fig. 3.2. B: MALDI-MS/MS spectrum of peak m/z 1088.4 (134-SKETSSPGT-142) resulting from subdigestion of the m/z 2820.1 peak.

In Figure 3.2, we also see the peak at m/z 2900.0, corresponding to the same sequence of the 2820.1 ion discussed above, except that this peptide is quadruply phosphorylated. The intensity of this peak was fairly low, leaving us unable to obtain good quality spectra from this ion even with the high amount of starting materials we were working with. This problem was further compounded by the fact that MALDI-MS/MS sensitivity as well as fragmentation efficiency decrease for high mass ions. However, as shown in Figure 3.5, after Asp-N sub-digestion of the OT-IMAC eluate, in addition to the 1088.4 peak discussed above, we also observed a peak at m/z 1221.5, corresponding to the peptide with sequence 126-DDGGIMMRSK-135. This was a fortuitous and unexpected find, since before sub-digestion with Asp-N, the original tryptic peptide would have the sequence 119-VAEEDEDDDGGIMMRSK-135—but was not observed in the original MALDI-MS spectrum. Perhaps something about the cluster of acidic E/D residues makes it difficult for this peptide to ionize in MALDI.

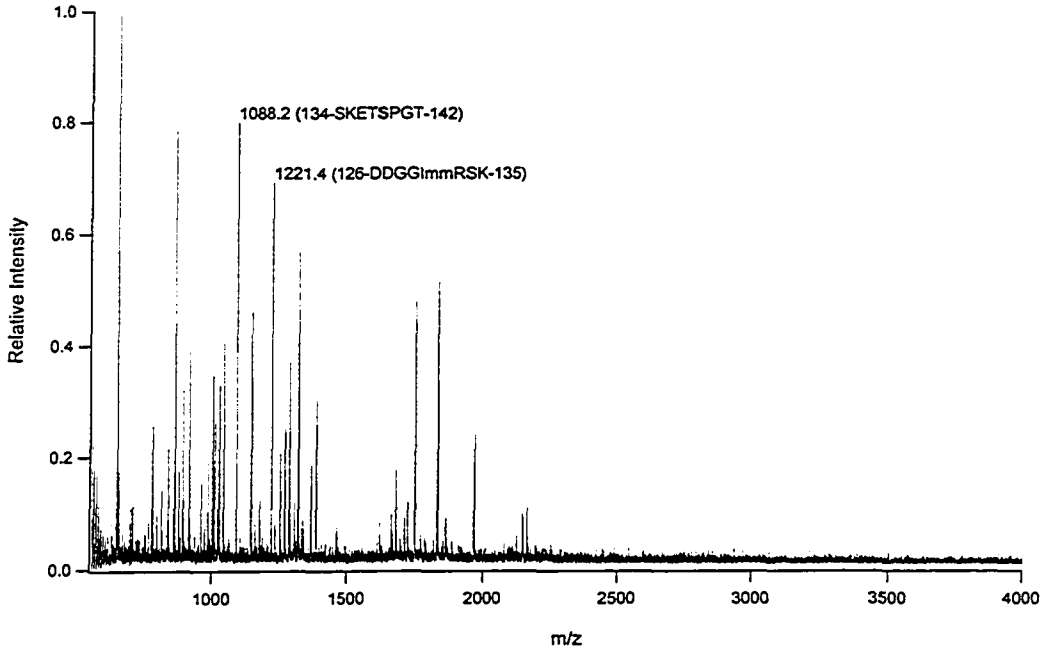

Figure3.5. MALDI-MS spectrum of the Asp-N sub-digested eluate analyzed in Fig. 3.2.

Fig.3.6 shows the MALDI-MS/MS spectrum obtained from the m/z 1221.5 ion. Assignment of the site of phosphorylation in this peptide is trivial, since it only contains one serine residue. However, had the situation been any different we would have had trouble assigning the site, since the quality of the MS/MS spectrum is not very good. Only the y_5-H_3PO_4-$2NH_3$ and y_5-H_3PO_4-$2NH_3$-$SOCH_4$ ions point to the presence of phosphate groups on this peptide (other than its MH^+ mass); however it's presence allows us to infer that the fourth site of phosphorylation on the m/z 2900.0 ion is located at Ser-134, since there is an overlap of 2 amino acids between the m/z 1221.5 ion and the m/z 2900.0 ion. For clarity this overlap is shown below.

134-SKETSSPGTDDVFTPAPSDSPSSQR-158
126-DDGGIMMRSK-135

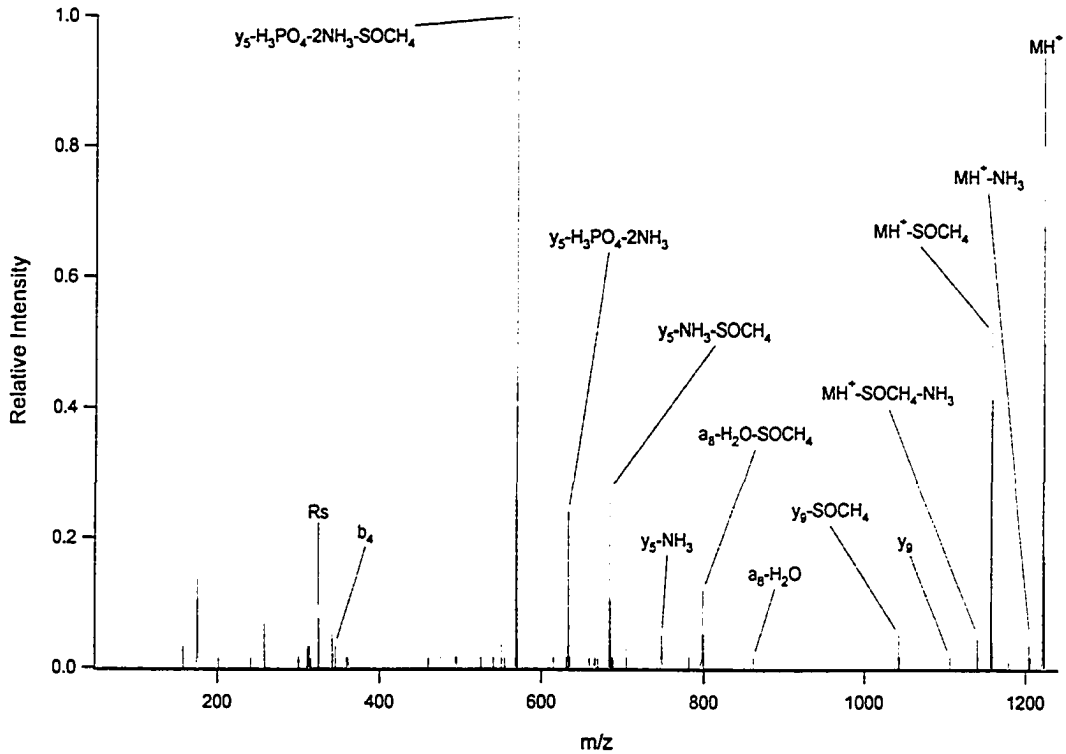

Figure 3.6. MALDI-MS/MS spectrum of the 1221.7 ion obtained after Asp-N subdigestion of the OT-IMAC elutate of His182.

From the foregoing experiments and assignments of MS/MS spectra, we conclude that 5 phosphorylation sites were successfully mapped in the His-182 Na^+/H^+ exchanger protein. These sites are Ser-61, Ser-134, Ser-138, Thr-147, and Ser-153. Figure 3.7 shows the total sequence coverage obtained from all peptides detected by MALDI-MS in the His182 protein. The sequence coverage represents 78% of the total protein, as 148 out of a total 189 amino acids were detected as peptides. Hence, 22% of the sequence remains undetected; therefore, there are 5 additional potential phosphorylation sites which have not been accounted for. In addition, detection of peptides without a phosphate group but having serines and/or threonines in their sequences does not necessarily mean that these are not phosphorylated; they could be

phosphorylated very inefficiently by the kinase, ending up with very low stoichiometry, although this is unlikely since a scenario of this type is more likely to play out *in vivo*. Achieving 100% sequence coverage however, still remains a challenge for any method relying on MS technologies for mapping of post-translational modifications. Future experiments on this biologically important protein in Prof. L. Fliegel's laboratory involve site directed mutagenesis of the identified phosphorylation sites in the full-length Na+/H+ exchanger protein. Mutation of the phosphorylated residues to nonphosphorylatable amino acids of equivalent size (ex./Ser to Ala, Thr to Val) and subsequent expression of the protein in a Na^+/H^+ exchanger deficient cells will then allow the biochemists to study the effects of its phosphorylation on regulation and activity of it's metabolic pathways.

1	11	21	31
MILRNNLQKT	**RQRLRSYNRH**	**TLVADPYEEA**	**WNQMLLRRQK**
41	51	61	71
ARQLEQKINN	**YLTVPAHKLD**	**sPTMSRARIG**	**SDPLAYEPKE**
81	91	101	111
DLPVITIDPA	SPQSPESVDL	VNEELKGKVL	**GLSRDPAKVA**
121	**131**	**141**	**151**
EEDEDDDGGI	**MMRsKETsSP**	**GTDDVFtPAP**	**SDsPSSQRIQ**
161	**171**	**181**	
RCLSDPGPHP	**EPGEGEPFFP**	**KGQHHHHHH**	

Figure 3.7. Total sequence coverage of the herein analyzed His182 Na^+/H^+ exchanger protein. Sequences in bold print were detected by MALDI-MS. Serines and threonines determined to be present as phosphoserine and phosphothreonine are shown in lower case.

3.3.2 Sensitivity Study on His-182 Using OTC-IMAC

In order to gauge the sensitivity of the capillary version of the open tubular IMAC on a real-world sample, we performed dilutions of His182 and separated the resulting His182 samples by SDS-PAGE. Figure 3.8 shows the separation obtained. His182 is a fusion protein with a histidine tag expressed in *E. coli,* which allows it to be purified using a commercial Ni-NTA IMAC beads; these IMAC beads show specificity towards binding of the polyhistidine motif. In looking at the gel, we can see that this purification strategy is quite good, with only minor non-His182 components being visible at higher concentrations (lanes 1 and 2). The molecular weight of His182 is 21325 Da without the presence phosphate groups, and as we see in the gel picture, the protein splits into a number of bands based on varying degrees of phosphorylation. Bands labeled A-D in Figure 3.8 were excised, tryptically digested, and then analyzed both with and without OTC-IMAC enrichment; furthermore, the derivatization with MeOH/HCl discussed in Chapter 2 was performed to study the usefulness of this approach for the analysis of this particular sample.

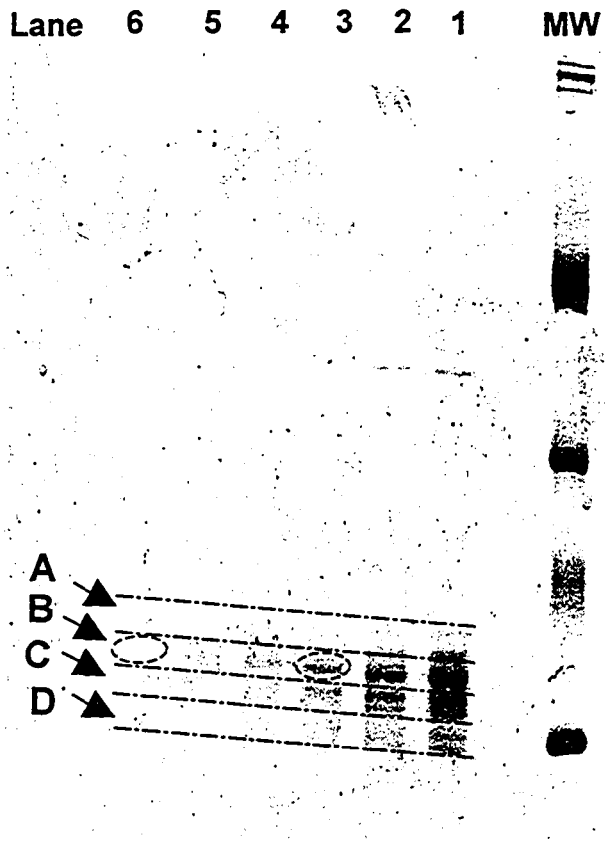

Figure 3.8. Dilutions of His182 separated on a 12% SDS-PAGE gel stained with Coomasie-blue dye. Lanes 1 through 6 contain 2.0 μg, 1.0 μg, 0.4 μg, 0.3 μg, 0.2 μg, and 0.1 μg total protein loads, respectively. Bands A – D were excised and analysed by MALDI-MS and OTC-IMAC.

3.3.3. Analysis of 0.5 μg In-Gel Digested His182: To Esterify or not to Esterify?

Having digested the bands from lane 2 shown in Figure 3.8 with trypsin, peptides were extracted from the gel, desalted by way of a C_{18} pipette tip, then esterified with the MeOH/HCl reagent described in Chapter 2.

Firstly, we noticed a difference between this His182 sample, and the one used previously for sequencing of the phosphorylation sites: phosphopeptides could now be seen in the spectra even prior to IMAC enrichment, albeit among a clutter of other non-phosphorylated peptides. The exact reason for this is unclear, perhaps owing to the longer digestion times used for this analysis, perhaps due to a difference in the quality and stoichiometry of the phosphorylation reaction itself. At any rate, with the new batch of His182, we observed a much less pronounced ion-suppresion effect acting on the phosphopeptides. Figure 3.9 shows a comparison of the analysis both with and without OTC-IMAC enrichment including a contrast between positive and negative-ion MALDI-MS. As is evident from this figure, OTC-IMAC enrichment successfully eliminates most non-phosphorylated peptides observed in panel A and further allows us to observe 4 other ones. As discussed in Chapter 2 for the case of α-casein, application of the negative mode for the esterified digest should in theory allow us to selectively ionize phosphopeptides provided they are present at a concentration where they can still be observed in positive ion mode. Indeed, this was found to be the case for this sample as well, as shown in panel B of Figure 3.9.

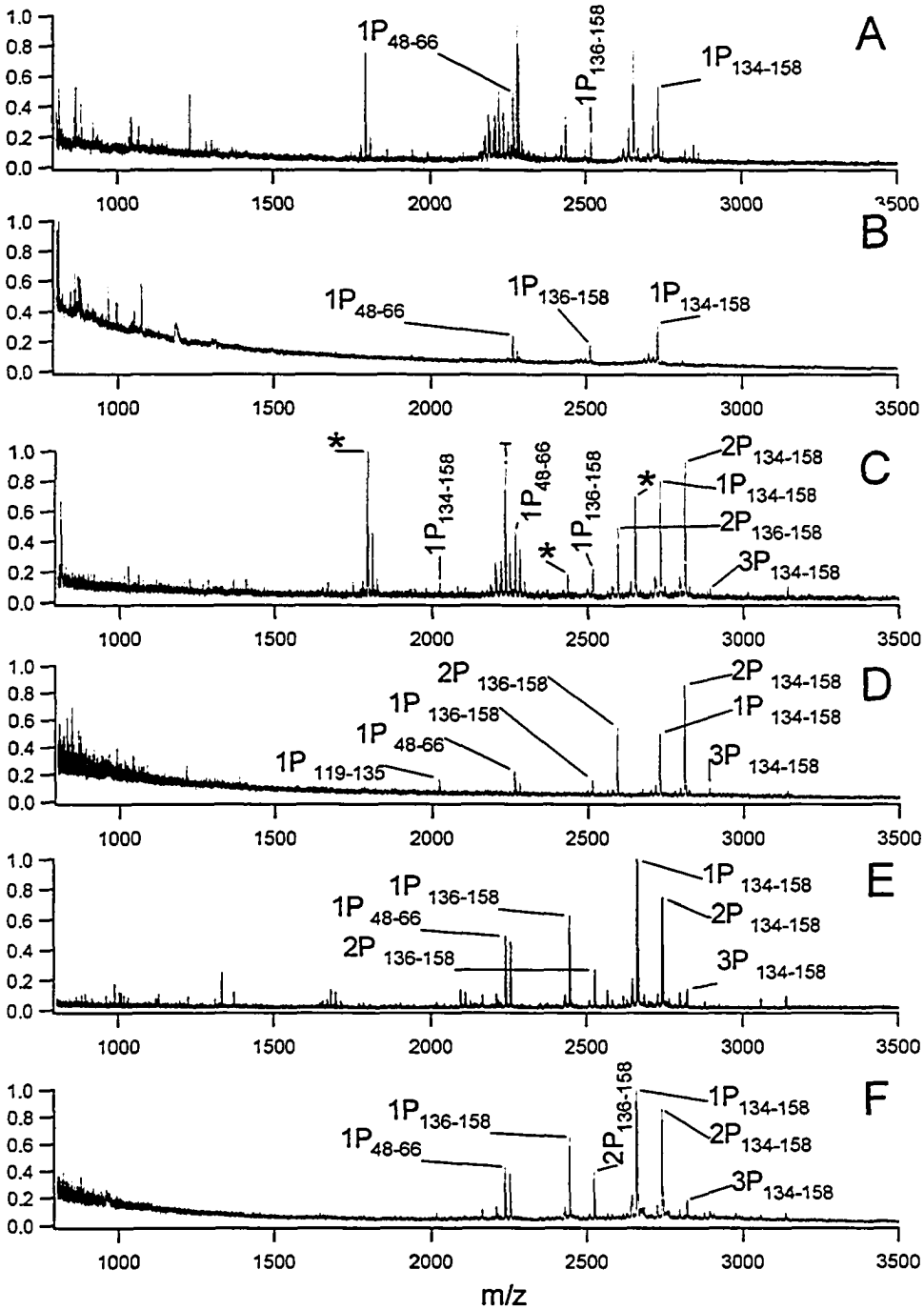

Figure 3.9. Comparison of OTC-IMAC enriched and original His182 digests of band B from Fig. 3.2.1. A: Esterified; no IMAC enrichment, positive-ion MALDI-MS; B: Esterified; no IMAC enrichment, negative-ion MALDI-MS; C: Esterified; OTC-IMAC enriched, positive-ion MALDI-MS D: esterified; OTC-IMAC enriched, negative-ion MALDI-MS.; E: OTC-IMAC enriched; positive-ion MALDI-MS; F: OTC-IMAC enriched, negative-ion MALDI-MS.

In comparing the esterified digest without OTC-IMAC enrichment in positive and negative ion mode MALDI-MS (Fig.3.9 A and B), it is evident that application of the negative ion mode eliminates the great majority of non-phosphorylated peptides detected in the positive ion mode. As shown in panel C, application of OTC-IMAC enrichment allows us to further concentrate phosphopeptides and detect them with greater sensitivity. Further analysis of panel C also shows that even with extensive washing, some esterified but non-phosphorylated peptides are still being retained on the column, presumably through hydrophobic interactions, as was discussed for α-casein in chapter 2. Again, as was the case with α–casein, application of the negative-ion mode to the OTC-IMAC eluate then allows us to further clean-up the spectrum by eliminating the now inefficiently ionizing non-phosphorylated peptides (panels C and D).

We also wanted to compare the above results with those for the underivatized sample, so see whether there is any significant advantage to esterifying carboxylate groups for this type of analysis; significant here meaning that it would be worth the extra time and effort and cost required for the analysis of the esterified digests. This is shown in Figure 3.9 E and F. It is evident that non-specific binding to IMAC is not a big concern, since the majority of peaks seen in panel C are indeed phosphorylated. In fact, if one compares panels C and E, it may appear that esterification of acidic groups makes at least some non-phohorylated peptides even more

'sticky' and retainable. When figure 3.9 is analyzed further, one will also see that following derivatization, we are also able to detect the 119-VAEEDEDDDGGIMMRSK-135 peptide at m/z 2022.8, which was previously undetected (undetectable?) without the esterification chemistry. This is the same peptide discussed on pg. 93-94; It was only detected after Asp-N sub-digestion of the tryptic digest enriched with OT-IMAC. The m/z of this peptide corresponds to the sequence shown above, with one phosphate group, and only 1 ester. Since there are a total of 8 acidic groups on this peptide, including the C-terminus, we would expect to observe this ion at m/z 2112.8 had it been fully esterified. Since there is no indication of incomplete esterification from any of the other peptides in the spectrum, we infer that most likely it is the long cluster of acid D/E residues and/or folding pattern of this peptide which prevents it from being fully derivatized; furthermore, since following sub-digestion of the said peptide into the smaller fragment with sequence 126-DDGGIMMRSK-135 allowed it to be detected in the other analysis (pg.98), confirms our suspicion that it is in fact this long cluster of acidic residues which prevented it from ionizing efficiently in MALDI-MS. Exactly why esterification of only 1 acidic residue would make it more amenable to IMAC enrichment and/or detection is not clear. However, we obtained an MS/MS spectrum of this peptide, shown in Fig. 3.10, to further confirm the phosphorylation site assignment made previously.

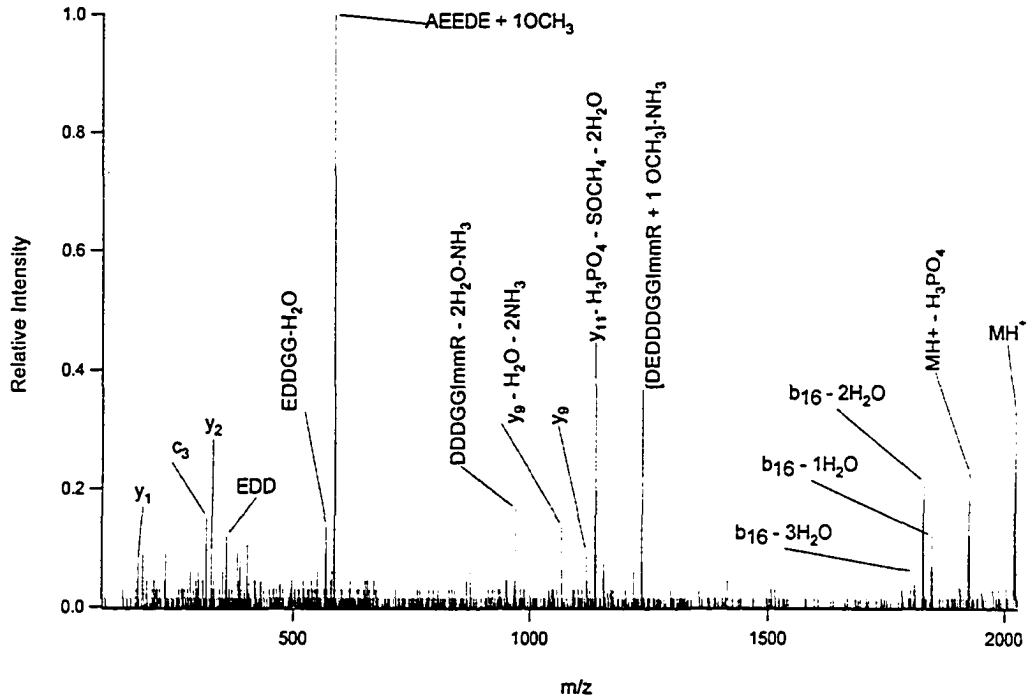

Figure 3.10. MS/MS Spectrum of the 119-VAEEDEDDDGGIMMRSK-135 peptide with mass 2022.8 Da, corresponding to it having 1 methyl esterification site.

Had it not been for the single peptide discussed above, we would conclude with confidence that the herein esterification chemistry is of little practical use for analysis of single proteins. As reported by Brill *et al.*[7], esterification becomes more useful in analysis and IMAC enrichment of phosphopeptides en masse, as in phosphoproteome analysis, for example.

Taking advantage of the enhanced detection sensitivity of OTC-IMAC, we also show that enrichment of phosphopeptides from trace-level digests becomes possible for a real-world sample. Fig. 3.11 shows the OTC-IMAC enriched phosphopeptides obtained from ~25-50 ng (1-2 pmol) of His182. We can see that even at this trace-level of protein content, which is below the detection limit of the dye used for gel staining,

phosphopeptide enrichment is still feasible. With the esterification chemistry, 3 phosphopeptides were detected; without esterification, a total of 4 were seen. The total number of phosphopeptides detected naturally decreases compared to analysis at higher levels, and obtaining meaningful MS/MS data is impossible; however, we're still very pleased to have been able to detect them at all at such low levels.

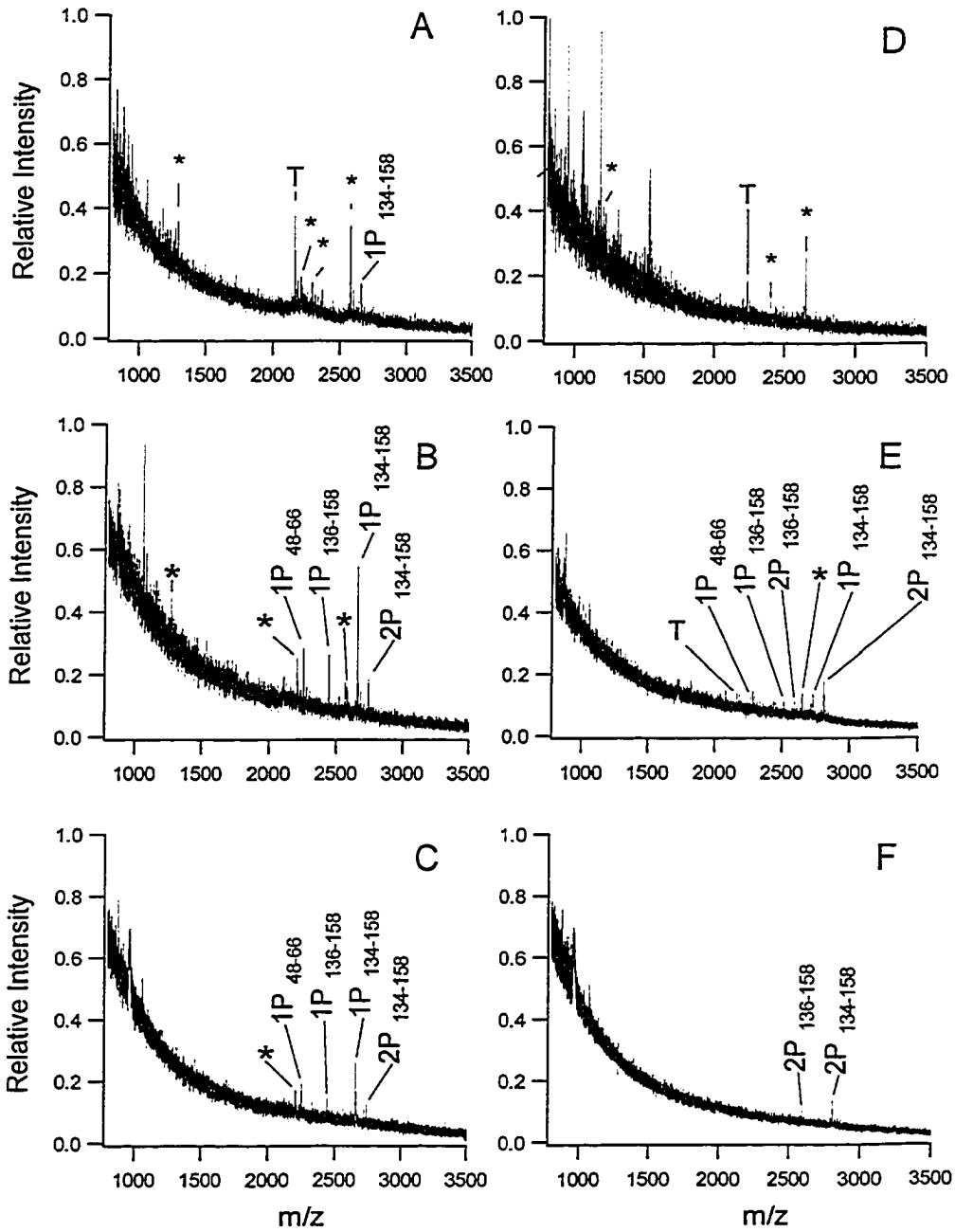

Figure 3.11. OTC-IMAC enrichment of trace-level peptides from His182. Band B from lane 6 in Fig.3.2.1, corresponding to ~25-50 ng of protein, was used for this analysis. A: No IMAC enrichment; positive-ion mode B: OTC-IMAC enrichment, positive ion mode. C: OTC-IMAC enrichment, negative-ion mode; D: No IMAC enrichment, with esterification; positive-ion mode E: OTC-IMAC enrichment, with esterification, positive-ion mode; F: OTC-IMAC enrichment, with esterification, negative-ion mode.

3.3.4. Conclusions

OT-IMAC IMAC was used to enrich phosphopeptides from The Na^+/H^+ exchanger protein His182, which was separated by SDS-PAGE. A total of 5 phosphorylation sites were mapped, they were: Ser-61, Ser-134, Ser-138, Thr-147, and Ser-153. 78% of the protein's sequence was detected by MALDI-MS following enzymatic digestion. It was shown that OTC-IMAC can detect phosphopeptides from a real sample separated by SDS-PAGE at levels lower than the detection limit of Coomasie blue staining. Much like the case for the α-casein data presented in Chapter 2, selectivity of silica-based IMAC is quite good for enrichment of phosphopeptides from a single protein, and we find no great advantage in applying esterification chemistry to eliminate non-specific binding; however, we found the esterification to be an alternative way to selectively ionize phosphopeptides in negative-ion MALDI-MS.

3.3.5. References Cited

1. Putney, L.K., Denker, S.P., Barber, D.L., **2002**, *Annu. Rev. Pharmacol. Toxicol.*, 42, 527-552.
2. Moor, A.N., Fliegel, L., **1999**, *J. Biol. Chem.*, 274, 22985-22992.
3. Wakabayashi, S., Bertrand, B., Shigekawa, M., Fafournoux, P., Pouyssegur, J., **1994**, *J. Biol. Chem.*, 269, 5583-5588.

4. Li, X., Liu, Y.S., Kay, C.M., Muller-Esterl, W., Fliegel, L., **2003**, *Biochemistry*, 42, 7448-7456.

5. Fliegel, L., Dyck, J.R.B., Wang, H., Fong, C., Haworth, R.S., **1993**, *Biochemistry*, 125, 137-143.

6. Mann, M., Jensen, O.N., **2003**, *Nature Biotech.*, 21, 255-261.

7. Brill, L.M., Salomon, A.R., Ficarro, S.B., Mukherji, M., Stettler-Gill, M., Peters, E.C., **2004**, *Anal. Chem.*, 76, 2763-2772.

Chapter 4

Conclusions and Future Work

OT-IMAC and OTC-IMAC enrichment is an effective way to enrich phosphopeptides from enzymatically digested proteins. Affinity capture and preconcentration is achieved, and most non-phosphorylated peptides are eliminated, thereby overcoming any ion-suppression effects in MALDI-MS.

In Chapter 2, we show that OT and OTC-IMAC have a sensitivity advantage over the popular IMAC product from Millipore. The sensitivity is highest with the capillary IMAC device and DHB MALDI matrix; using this approach, we can enrich phosphopeptides from 10 ng (400 fmol) of α-casein separated by SDS-PAGE and digested with trypsin in-gel, and 0.2 ng (20 fmol) α-casein digested with trypsin in-solution. Esterification chemistry developed by Ficarro *et. al.* has been scaled down to handle sub-pmol amount of protein separated by SDS-PAGE; however, we did not observe a great advantage in using this approach to increase the selectivity of IMAC for single proteins digested with trypsin. Application of the negative-ion mode in MALDI-MS, when combined with esterification chemistry, has been shown to allow for much improved ionization selectivity of derivatized phosphopeptides.

In Chapter 3, OT-IMAC combined with double enzyme digestion and MALDI-MS/MS was applied towards mapping of phosphorylation

sites in the Na^+/H^+ exchanger protein His182. A total of 5 phosphorylation sites were unambiguously mapped to Ser-61, Ser-134, Ser-138, Thr-147, and Ser-153. Only 22% of the protein's sequence was left undetected, containing 5 potential phosphorylation sites. It was shown that OTC-IMAC can detect phosphopeptides His182 separated by SDS-PAGE at levels lower than the detection limit of Coomasie-blue staining. We found the selectivity of silica-based IMAC to be quite good for enrichment of phosphopeptides from a single protein, and found no great advantage in applying esterification chemistry to eliminate non-specific binding; however, we found the esterification to be an alternative way to selectively ionize phosphopeptides in negative-ion MALDI-MS.

While my thesis research has been involved in developing a facile technique for phosphoprotein characterization based on manual sample handling, it provides the basis for future work in developing a more automated sample handling system. To this end, the manual spotting technique used with OTC-IMAC has been extended into an automated capillary spotter for HPLC effluent spotting directly onto a MALDI target for MS analysis which is a subject of another graduate student, Bryce Young's thesis work in our lab. Combining on-line IMAC enrichment with LC separation is a very attractive idea, especially for applications like phosphoproteomics of whole-cell extracts, where literally hundreds of phosphopeptides would be expected to bind the affinity column. For such an application, one of the limitations of the herein described open-tubular IMAC devices is the binding capacity, where the IMAC surface would be

expected to reach saturation levels. For this purpose, we've also synthesized IMAC directly on glass beads, which could then be used to pack an LC IMAC-precolumn for automated on-line phosphopeptide enrichment. Recently, it has been reported that phosphoric acid enhances MALDI-MS detection of phosphopeptides[1], and this could further be used to fine-tune our methods. We envision that an automated sample handling system tailored to single phosphoprotein analysis using open tubular IMAC as well as complicated phosphoproteins analysis using packed IMAC column will be very useful for characterization of phosphoprotein or phosphoproteome. For now, the technique described in this thesis should be useful to handle gel-separated phosphoproteins, albeit with manual operation.

[1] Stensballe, A., Jensen, O.N., **2004**, *Rapid Commun. Mass Spectrom.*, 18, 1721-1730