NEW STRATEGIES IN FUNCTIONAL PROTEOMICS

Claudia Cirulli

Dottorato in Scienze Biotecnologiche – XIX ciclo
Indirizzo Biotecnologie Molecolari
Università di Napoli Federico II

Dottorato in Scienze Biotecnologiche – XIX ciclo
Indirizzo Biotecnologie Molecolari
Università di Napoli Federico II

NEW STRATEGIES IN FUNCTIONAL PROTEOMICS

Claudia Cirulli

Dottoranda: Claudia Cirulli

Relatore: Prof. Gennaro Marino

Coordinatore: Prof. Gennaro Marino

Nessuno ha un amore più
grande di questo: dare
la vita per i propri amici.

Giov15, 13

INDICE

Riassunto

Il termine “proteomica” è utilizzato per indicare la procedura di identificazione di un elevato numero di proteine in miscele molto complesse provenienti da organelli cellulari o, addirittura, interi microrganismi. I moderni studi sul proteoma possono essere divisi essenzialmente in due settori principali: la proteomica di espressione che è rivolta alla definizione qualitativa e quantitativa dei livelli di espressione delle proteine, e la proteomica funzionale il cui obiettivo è quello di definire il ruolo delle proteine nella cellula e descrivere le vie di trasduzione del segnale. E’ noto che nelle cellule molti processi sono governati non solo dall’abbondanza relativa delle proteine ma anche da regolazioni transienti dell’attività, dell’associazione e localizzazione di proteine e complessi proteici. Queste proprietà sono sottilmente controllate da una serie di modifiche post-traduzionali eventualmente reversibili (1).

Una delle più importanti modifiche è senza dubbio la fosforilazione, che influenza l’attività delle proteine e i punti critici di numerosi processi cellulari.

La fosforilazione si comporta come “interruttore” molecolare o modulatore dell’attività e della localizzazione delle proteine e regola l’assemblaggio e la dissociazione di complessi macromolecolari. Per capire come funzionino questi meccanismi è necessario identificare la proteina coinvolta in un particolare processo cellulare e lo specifico residuo amminoacidico modificato. Questa informazione non può essere determinata dalla sequenza del gene (2). Infatti, dal codice genetico è possibile risalire alla sequenza primaria di una proteina ma non è possibile stabilire e capire il tipo di modifica a cui sarà sottoposta e in quale momento della vita cellulare verrà modificata e/o quale amminoacido verrà coinvolto nella modifica.

Tuttavia, la rilevazione della fosforilazione incontra varie difficoltà: la presenza in una proteina di più possibili siti di fosforilazione determina una particolare eterogeneità molecolare, a causa della differente specificità delle chinasi, (3). Un’ulteriore difficoltà è dovuta al fatto che la fosforilazione è sub-stechiometrica poiché solo una parte della popolazione molecolare di proteina substrato viene modificata in un particolare momento della vita cellulare, in seguito ad un segnale assai spesso transiente.

Il metodo più comune per l'analisi di fosfoproteine comporta l'uso di traccianti radioattivi *in vivo* e/o *in vitro*. Un approccio alternativo, meno sensibile, consiste nell’uso di coloranti fluorescenti. Il Pro-Q Diamond, appartiene a questa classe di reattivi, che permettono di rivelare selettivamente le fosfoproteine separate mediante elettroforesi SDS-PAGE. Altri approcci si basano sull’ arricchimento preliminare delle fosfoproteine, utilizzando anticorpi specifici come nel caso delle proteine con residui di fosfotirosina; altri ancora sfruttano la capacità del gruppo fosfato di coordinarsi a metalli come il gallio o ferro (IMAC) (4). Tuttavia questa strategia presenta lo svantaggio di avere specificità variabile a causa dell’affinità dei cationi per i gruppi acidi, carichi negativamente(5). Una più fine identificazione dei siti di fosforilazione è in genere condotta attraverso frazionamento HPLC (high pressure liquid cromatografy) delle miscele peptidiche seguito dall’analisi mediante spettrometria di massa tandem delle singole frazioni (6). Utilizzando sistemi cromatografici nanoHPLC e spettrometri di massa “ibridi” è possibile raggiungere una maggiore sensibilità rispetto alle strumentazioni classiche.

Un progetto inteso alla caratterizzazione dell’insieme di proteine che sono modificate *in vivo* mediante fosforilazione è alla base della comprensione dei

meccanismi con cui le proteine sono attivate o inibite all'interno della cascata della trasduzione del segnale (fosfoproteoma).

La prima parte del lavoro di tesi è stata rivolta all'elaborazione di una strategia che ha reso possibile l'identificazione e l'analisi quantitativa di proteine fosforilate in corrispondenza dei residui di P-Ser e P-Thr. La messa a punto di questa strategia ha previsto l'utilizzo di una proteina modello fosforilata in serina: l'alfa caseina. I peptidi O-fosforilati sono stati sottoposti ad una reazione di beta eliminazione del gruppo fosfato, ottimizzando le condizioni già riportate in letteratura. I residui di deidroalanina, ΔSer, e acido deidro-ammino 2,butirrico, ΔThr sono stati modificati per addizione di una delle due funzionalità tioliche del ditiotreitolo (DTT) sul gruppo vinilico generato.

L'introduzione di un gruppo tiolico in corrispondenza del sito iniziale di fosforilazione ha permesso, mediante un passaggio di purificazione con una resina attivata a gruppi tiolici, l'isolamento selettivo dei soli peptidi fosforilati. Questi ultimi sono stati identificati mediante spettrometria di massa MALDI ed ESI. La spettrometria di massa tandem ha permesso di ottenere le informazioni necessarie sulla sequenza dei peptidi e l'esatta localizzazione del sito di fosforilazione. Il problema dell'analisi quantitativa è stato affrontato utilizzando l'approccio dei reattivi della classe di tipo ICAT (7). L'analisi quantitativa relativa è stata realizzata utilizzando uno standard interno ottenuto mediante l'uso del reattivo, il DTT, completamente deuterato. L'analisi quantitativa dei peptidi è stata condotta misurando le intensità relative delle coppie di peptidi, ottenuti per analisi MALDI/MS e LC-ESI/MS, aventi la stessa sequenza, ma marcati con le forme leggere e pesanti del reattivo, e quindi dotati di un peso molecolare differente di 6 Da in base al numero di atomi di deuterio presenti nel reattivo utilizzato. Per verificare la potenzialità di questa strategia nel determinare la quantità relativa di proteine fosforilate presenti in due differenti stati cellulari sono stati preparati campioni contenenti concentrazioni delle due proteine derivatizzate con DTT (D0) e DTT (D6), in rapporti molari pari a 1:1, 2:1, 4:1. I dati ottenuti dall'analisi di massa hanno confermato l'applicabilità di questa strategia per questo tipo di analisi (8).

Nella seconda parte del lavoro di tesi sono state ottimizzate le condizioni sperimentali per il protocollo di arricchimento di fosfopeptidi in miscele complesse mediante *ion metal affinity cromatography* (IMAC). Inoltre sono stati messi a punto i parametri di acquisizione per analisi selettiva di fosfopeptidi mediante tecniche di spettrometria di massa (*precursor ion scan, neutral loss and multiple reaction monitorino MRM*). Per mettere a punto la strategia e verificarne l' efficienza è stata utilizzata una proteina modello, l'α-caseina. Lo ione Fe(III) è stato scelto per effettuare la cromatografia IMAC. Il risultato di questo tipo di cromatografia è influenzato da vari fattori quali lo ione metallico, il gruppo chelante, le condizioni di legame e di lavaggio nonché quelle di eluizione. Per questo motivo, il primo passo è stato la scelta del gruppo chelante lo ione metallico più appropriato. L' α-caseina idrolizzata con tripsina è stata sottoposta alla cromatografia di affinità utilizzando sia la resina NTA che IDA aventi rispettivamente come gruppo chelante l'acido nitrilo-triacetico e l'acido immino-diacetico. Questa analisi è stata ripetuta più volte e i ritenuti provenienti dalle due cromatografie sono stati sottoposti all'analisi mediante spettrometria di massa MALDI. La resina IDA è stata scelta sulla base della maggiore presenza di specie fosforilate rispetto a quelle non fosforilate negli spettri MALDI. A questo punto prendendo come punto di partenza le condizioni riportate in letteratura sono state apportate delle modifiche che hanno permesso di isolare tutti i peptidi fosforilati dell'α-caseina. Inoltre è stato possibile riscontrare negli spettri la presenza di alcuni

peptidi non fosforilati ricchi di residui amminoacidici acidi. Infatti, questo tipo di strategia possiede la capacità di interagire preferenzialmente con la carica negativa del gruppo fosfato, ma presenta lo svantaggio di avere specificità variabile per i gruppi acidi, carichi negativamente Per ovviare a questo inconveniente è stata effettuata una reazione di esterificazione dei gruppi carbossilici delle miscele peptidiche, come proposto da Ficarro et. al (9), prima di sottoporle al passaggio di arricchimento. L'analisi MALDI-MS delle miscele peptidiche esterificate e arricchite ha sottolineato la presenza di tutti i peptidi fosforilati dell'alfa caseina, incluso il peptide penta fosforilato ed un decremento di peptidi aspecifici. È stato riscontrato, inoltre, che la reazione di esterificazione è solo parziale, generando eterogeneità molecolare. Questo fattore diminuisce la sensibilità nella rivelazione dei peptidi fosforilati. Per questo motivo, per minimizzare la presenza di peptidi contenenti residui acidi durante la cromatografia IMAC il pH della soluzione contenente campione è stato portato a 2.5. A questo valore di pH, i gruppi carbossilici presenti nella proteina sono protonati mentre il gruppo fosfato, è deprotonato in base ai differenti valori del pka dei gruppi carbossilici rispetto a quello dell'acido fosforico (pka_1 1,8). Per verificare la validità di questa strategia per l'isolamento di specie fosforilate in miscele complesse, l'α-caseina è stata introdotta in un campione di siero. Quest'ultimo rappresenta un campione molto complesso le cui proteine più abbondanti sono l'albumina, le immunoglobuline di classe G, la trasferrina etc. L'analisi MALDI ha permesso di confermare la presenza delle specie fosforilate già precedentemente attribuite e una contemporanea diminuzione di peptidi aspecifici. Le miscele peptidiche arricchite di peptidi fosforilati sono state utilizzate per ottimizzare i parametri di acquisizione delle scansioni di *precursor ion* di *neutral loss* e di *MRM.* Infatti, utilizzando uno spettrometro di massa ibrido a trappola ionica lineare è stato possibile effettuare questo tipo di scansioni. Ad esempio, nel caso della *precursor ion scan* è stato possibile rivelare selettivamente i peptidi fosforilati, sfruttando la loro peculiare caratteristica di generare un particolare frammento a -79 m/z, dovuto alla perdita del gruppo fosfato. La strategia messa a punto è stata utilizzata per identificare i peptidi liberi presenti in fluidi biologici come il siero, la saliva e le urine. Lo studio di questi peptidi e la caratterizzazione del loro stato di modifica post tradizionale sta diventando sempre più interessante a causa dell'informazione che possono fornire in merito ad alcune patologie. Essi possono rappresentare dei potenziali biomarker per diagnosticare numerose malattie (10). Questo studio ha permesso di identificare 13 diversi peptidi fosforilati liberi e 7 differenti siti di fosforilazione.

L'ultima parte di questo lavoro è stata impegnata nello sviluppo di una procedura di marcatura chimica selettiva di residui di fosfoSer/Thr con cloruro di dansile (5-dimetilammino-1-naftalene sulfonico cloruro) usata in combinazione a tecniche avanzate di spettrometria di massa (MS^3).

L'agente derivatizzante, dansil cloruro, è comunemente usato nell'analisi di proteine possedendo elevata reattività nei confronti di gruppi amminici primari e secondari. La sua capacità di fluorescenza consente di aumentare il limite di rivelabilità degli analiti derivatizzati. Fin dal 1968, sono state notate caratteristiche frammentazioni dei dansil-derivati, capaci di produrre ioni frammento a 170 m/z e 234 m/z (11). Queste frammentazioni possono essere sfruttate per l'analisi selettiva di modifiche post tradizionali. In particolare, possono essere utilizzati come ioni *reporter* in scansioni del tipo *precursor ion scan*. Nel laboratorio dove ho svolto il mio lavoro di dottorato sono state studiate le caratteristiche frammentazioni di dansil derivati in strumenti a trappola ionica lineare. In questi studi e stato notato che lo ione

frammento a 234 m/z, a seguito di un ulteriore evento di frammentazione (MS^3), è in grado di generare lo ione frammento a 170 m/z. In questo modo i dansil-peptidi possono essere rivelati selettivamente utilizzando analisi di tipo *precursor ion scan*, grazie al caratteristico ione frammento a 170 m/z. Un'ulteriore selezione degli ioni può essere realizzata sfruttando un'analisi MS^3 dello ione 234 m/z. Il protocollo è stato messo a punto su una miscela peptidica di α-caseina idrolizzata con tripsina. La strategia ha previsto la reazione di beta eliminazione dei gruppi fosfato dei residui di p-Ser e p-Thr e successiva derivatizzazione mediante addizione di Michael di N-(2-mercaptoetil) dansilamide (DANSH). Quest'ultima non disponibile commercialmente è stata sintetizzata a partire dal cloruro di dansile e dalla 2,2'-ditiobis(etilammina). La miscela di peptidi è stata analizzata mediante LC-MSMS. In questo caso l'utilizzo dell'analisi MS^3 ha consentito di ridurre notevolmente la complessità della miscela peptidica. In particolare sono state rilevate due specie associate ai peptidi monofosforilati dell'α-caseina (12).

In conclusione, nel corso del dottorato sono state elaborate tre differenti strategie di analisi integrate a tecniche avanzate di spettrometria di massa. Queste tecniche hanno permesso di studiare la fosforilazione e affrontare le diverse problematiche connesse a questo tipo di modifica post-traduzionale. È chiaro, che queste strategie possono trovare un'ampia applicazione in ambito proteomico, nella caratterizzazione dello stato di fosforilazione di vari sistemi biologici per la comprensione dei meccanismi molecolari alla base della vita cellulare; in ambito farmaceutico, nell'identificazione di nuovi biomarker per la diagnosi e la prognosi di numerose malattie.

Bibliografia

1. Turecek F, et al.(2002) J.mass spectrom. 37: 1-14.
2. Kjellstrom S., et al. (2004) Anal. Chem. 76:5109-5117.
3. Liu H., et al. (2004). Anal.Chem. 76:4223-4232.
4. Mann M., et al. (2002) Trends in biotec. Review 20: 261-267.
5. Tao He, et al. (2004). J. Am. Soc. Mass Spectrom. 15, 363-373.
6. Martijn, et al.(2004) Anal. Chem. 76: 3935-3943.
7. Tao He, et al. (2004). J. Am. Soc. Mass Spectrom. 15, 363-373.
8. Amoresano A., et al (2004). Eur J Mass Spectrom. 10, 401-412.
9. Ficarro SB, et. al. (2005). Rapid commun Mass Spectrom. 19, 57-71.
10. Lee HJ, et. al. (2006) Current opin chem. Biol. 10, 42-49.
11. Marino G., Buonocore V., Biochem J., 1968 110, 603.
12. Amoresano A., et. al. (2006). Rapid commun Mass Spectrom. 20,1400-1404.

Summary

The proteome is defined as the set of all expressed proteins in a cell, tissue or organism. Proteome analysis presents specialized analytical problems in two major areas: i) dynamic expression range, expression proteomics, which aims to measure up- and down-regulation of protein levels and (ii) diversity of protein expression (multiple protein forms), functional proteomics, aimed at the characterisation of cellular compartments, multi-protein complexes and signalling pathways.

One of the descriptors of a protein, which is amenable to proteomics technology, is the delineation of post-translational modifications (PTMs). The most important and abundant PTM that is used to modulate protein activity and propagate signals within cellular pathways and networks is the phosphorylation.

There are many classical techniques designed to determine whether a protein is phosphorylated.

This PhD thesis focuses on the development of new proteomic approaches based on classical biochemical methods coupled to new mass spectrometry methods to study the phosphorylation. These phosphoproteome studies aim at comprehensive analysing protein phosphorylation by identification of the phosphoproteins, exact localization of the residues that are phosphorylated, and preferably quantification of the phosphorylation.

Due to low stoichiometry, heterogeneity and low abundance, enrichment of phosphopeptides is an important step of this analysis.

The first section is focused on the development of new enrichment methods coupled to mass spectrometry. Thus, improved approach, based on simple chemical manipulations and mass spectrometric procedures, for the selective analysis of phosphoserine and phosphothreonine in protein mixtures, following conversion of the peptide phosphate moiety into DTT derivatives, is described. However the major aim of this work is devoted to the use of isotopically labelled DTT thus allowing a simple and direct quantitative MS analysis.

The second part aims to the optimization of the IMAC (Ion, Metal, Affinity, Chromatography) protocol coupled to iterative mass spectrometry based scanning techniques (neutral loss, precursor ion, multiple reactions monitoring) to improve detection of phosphorylated peptides.

A final part of the work is focused on the development of a strategy to study phosphorylation without preliminary enrichment but using the high performance of a novel hybrid mass spectrometer Linear Ion Trap.

To prove their wide enforceability these methodologies were set up applied first on a model phosphoprotein, bovine α-casein, and then applied to different biological systems.

I. Introduction

I.1 Proteomics Revolution

Proteomics — the analysis of genomic complements of proteins — has burst onto the scientific scene with stunning rapidity over the past few years, perhaps befitting a discipline that can enjoy the virtually instantaneous conversion of a genome sequence to a set of predicted proteins.

In fact, for almost two decades, major efforts have been directed at the polynucleotide level and, particularly, at the gene sequencing of a variety of different organisms.

Complete genome sequencing has been achieved for a wide variety of organisms, and efforts to sequence the complete human genome have resulted in completion of a first draft. Complete genomes for a steadily increasing number of organisms are now available (*http://www.ncbi.nlm.nih.gov/Entrez/Genome/main_genomes.html).*

While it is often conceptualized that one gene produces one protein, it is known that the expressed products of a single gene in reality represent a protein population that can contain large amounts of microheterogeneity. More than 100 modification types are recorded and additional ones are yet to be discovered (1). All modified forms from one protein can vary in abundance, activity or location inside a cell. Indeed, cellular proteins are not invariant products of genes, but are subject to a high degree of interdependent processing at the protein level that is a critical component of cellular function and regulation. Clearly, post-translational modification of proteins is an event with dramatic effects on the complexity of the proteome. In addition, ability of a proteome is its capacity to regulate dynamically protein expression in response to external and internal perturbations under developmental, physiological, pathological, pharmacological and aging conditions. In fact, in contrast to the static genome, where all information could in principle be obtained from the DNA of a single cell, the proteome is dynamic and highly dependent not only on the type of cell, but also on the state of the cell. This means that genome sequencing will not be able to define all of the actors (proteins) in the forms in which they actually exist in cells.

Proteome analysis presents specialized analytical problems in two major areas: i) dynamic expression range, **expression proteomics**, which aims to measure up- and down-regulation of protein levels (2) and (ii) diversity of protein expression (multiple protein forms), **functional proteomics**, aimed at the characterisation of cellular compartments, multi-protein complexes and signalling pathways (3).

The dynamic range problem can be overcome by either increased separation power or pre-fractionation to enrich for lower abundance proteins (4).
Overcoming the diversity of protein expression is more involved and represents a significant challenge for proteome (5).

It is becoming increasingly clear that the field of functional proteomics is to monitor and analyze the spatial and temporal properties of the molecular networks and fluxes involved in living cells and to identify the molecular species that participate in such networks via functional stimulation, perturbation, or isolation of these networks.

I.2 Proteomics methods: qualitative and quantitative proteome analysis

I.2.1 The first Generation Technology

The complexity of a proteome can far exceed the capacity of analytical systems. The whole proteome of any organism, in fact is too complicated to be analyzed in a simple one-step process and direct attempts for the entire proteome analysis normally lead to limited amount of information (6).

Classical proteomic approaches have relied upon separating whole cell lysates by 2D Gel Electrophoresis (2D-GE). Historically, 2D-GE has been the tool of choice to resolve complex protein mixtures and to detect differences in protein expression patterns between normal and diseased tissue.

High-resolution gel electrophoresis, of which 2D-GE is currently the most powerful protein separation method, was already used as an analytical tool in the late 70s (7). The 2D-GE consists in the separation of proteins by isoelectric point in one dimension and molecular weight in the second dimension. Despite these outstanding properties and widespread application, the systematic use of 2D-GE became an integral part of the proteomics 20 years ago. The protein spots were elettroblotted on the membrane and analyzed by Edman degradation to release N-terminal sequence. In fact, the Edman-based approach was generally slow because every protein spot or peptide peak had to be sequenced individually. In addition this identification method had low sensibility and there was not complete database of genome. Furthermore, proteins are often blocked at their N-terminal (acetylation, formylation or pyroglutamic formation).

2D-GE remained the highest resolution protein separation method available, but the ability to identify the observed proteins was always an extremely difficult problem. Although improvements in 2D gel technology had been realized since its introduction, three enabling technological advances have provided the basis for the foundation of the field of proteomics. The first advance was the introduction of large-scale nucleotide sequencing of both expressed sequence tags (ESTs) and, more recently, genomic DNA. The second was the development of mass spectrometers able to ionize and mass-analyze biological molecules and, more recently, the wide-spread introduction of mass spectrometers capable of data-dependent selection for fragmentation (MS/MS) (i.e., without the need for user intervention). The third was the development of computer algorithms able to match uninterpreted (or partially interpreted) MS/MS spectra with translations of the nucleotide sequence databases, thereby tying the first two technological advances together. Thus MS played a key role in the passage of 2-DE/image analysis to proteomics. In fact nowadays, mass spectrometry is considered as base technology for protein identification from 2D gels (8). However, mass spectrometric identification, although quite sensitive, is limited by the ability to reproducibly observe distinct spots on the 2-D gel electrophoresis.

Well-known drawbacks of the technique are limitations in the pI and molecular weight of proteins. In fact, specific classes of proteins have long been known to be excluded or under-represented in 2D gel patterns. These include very acidic or basic proteins, excessively large or small proteins and membrane proteins. Furthermore, difficult automation and reproducibility problems and low abundant proteins are either not detected at all due to the limited sensitivity achievable with commonly used dyes,

or they are masked by higher abundant comigrating proteins. Incremental improvements in 2DE technology, including more sensitive staining methods (10-11), large-format higher resolving gels (12) and sample fractionation prior to 2DE have alleviated, but not eliminated, these and other shortcomings of the 2DE/MS approach.

To address the problem of quantification, global protein expression analysis or quantitative proteomics has historically been carried out using 2D gels. The comparative 2D gel approach is typically employed to identify proteins that are up- or down-regulated in a disease state. In such studies, a reliable analysis of quantitative changes of protein spots is important (9). Furthermore, recently new programs for comparison of 2D-maps like algorithms based on characteristics of spot image on gel (PDQuest, Phoretix 2D and Melanie); or algorithms based on direct comparison of images by distribution of intensity (Z3 and MIR) are rising.

I.2.2 Protein Identification by mass spectrometry

To measure the mass of molecules, the test material must be charged (hence ionized) and desolvated (dry). The two most successful mechanisms for ionization of peptides and proteins are matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI). In MALDI the analyte of interest is embedded in a matrix that is dried and then volatilized in a vacuum under ultraviolet laser irradiation (13-14). Typically, the mass analyzer coupled with MALDI is a time-of-flight (TOF) mass analyzer that simply measures the elapsed time from acceleration of the charged (ionized) molecules through a field-free drift region. The other common ionization source is ESI, in which the analyte is sprayed from a fine needle at high voltage toward the inlet of the mass spectrometer at a lower voltage. The spray is typically either from a reversed phase HPLC (RP-HPLC) column or a nanospray device that is similar to a microinjection needle. During this process, the droplets containing analyte are dried and gain charge (ionize). The ions formed during this process are directed into the mass analyzer, which could be either a triple-quadrupole, an ion trap, a Fourier-transform ion cyclotron resonance (FT-ICR), or a hybrid quadrupole TOF (Qq-TOF) type (15-16-8) (Fig.1).

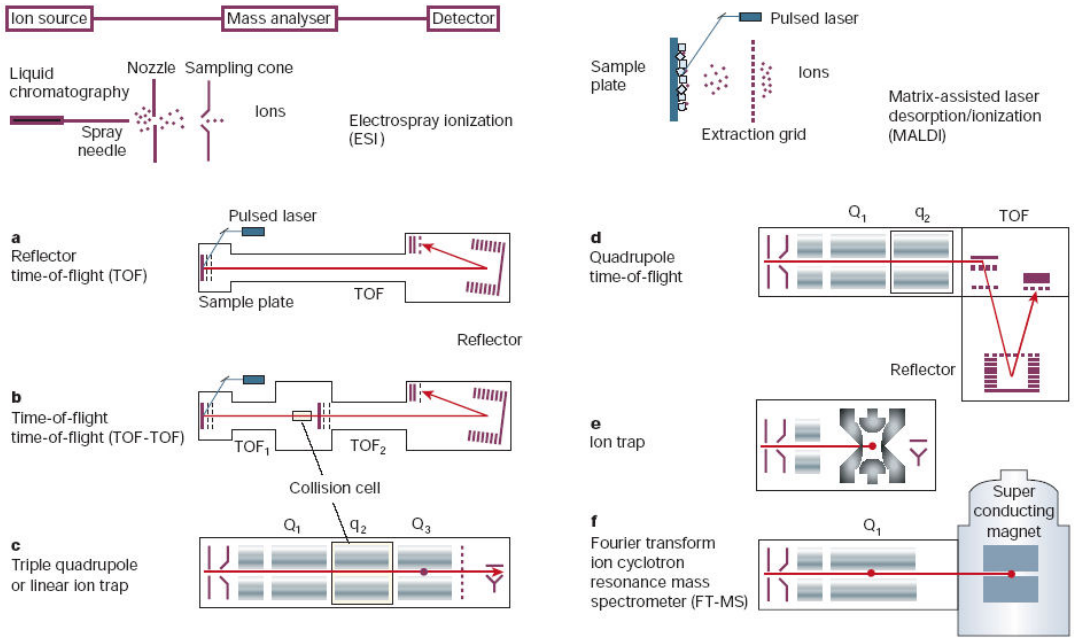


Figure 1. Mass spectrometers used in proteome research. The left and right upper panels depict the ionization and sample introduction process in electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI). The different instrumental configurations (a–f) are shown with their typical ion source. a, In reflector time-of-flight (TOF) instruments, the ions are accelerated to high kinetic energy and are separated along a flight tube as a result of their different velocities. The ions are turned around in a reflector, which compensates for slight differences in kinetic energy, and then impinge on a detector that amplifies and counts arriving ions. b, The TOF-TOF instrument incorporates a collision cell between two TOF sections. Ions of one mass-to-charge (m/ z) ratio are selected in the first TOF section, fragmented in the collision cell, and the masses of the fragments are separated in the second TOF section. c, Quadrupole mass spectrometers select by time-varying electric fields between four rods, which permit a stable trajectory only for ions of a particular desired m/ z. Again, ions of a particular m/z are selected in a first section (Q1), fragmented in a collision cell (q2), and the fragments separated in Q3. In the linear ion trap, ions are captured in a quadruple section, depicted by the red dot in Q3. They are then excited via resonant electric field and the fragments are scanned out, creating the tandem mass spectrum. d, The quadrupole TOF instrument combines the front part of a triple quadruple instrument with a reflector TOF section for measuring the mass of the ions. e, The (three-dimensional) ion trap captures the ions as in the case of the linear ion trap, fragments ions of a particular m/ z, and then scans out the fragments to generate the tandem mass spectrum. f, The FT-MS instrument also traps the ions, but does so with the help of strong magnetic fields. The figure shows the combination of FT-MS with the linear ion trap for efficient isolation, fragmentation and fragment detection in the FT-MS section.

Two mass spectrometric methods for rapid identification of proteins are now in widespread use. One of these methods uses the characteristic distribution of peptide masses obtained by chemical or enzymatic fragmentation of proteins (17). A number of computer programs are available for using the observed peptide masses to search gene sequence databases for proteins that fit the mass fingerprint (18-19-20-21). The other method makes use of sequence tags (22-23); i.e., partial amino acid sequence information. Searching of gene sequence databases is again used to identify the protein at the gene level.

Indeed, the typical proteomics experiment consists of six stages. In stage 1, the proteins to be analyzed are isolated from cell lysate or tissues by biochemical fractionation, 2D-GE separation or affinity selection. Therefore, proteins are degraded enzymatically to peptides in stage 2, usually by trypsin, leading to peptides with C-terminally protonated amino acids, providing an advantage in subsequent MS

analysis. In stage 3, proteins are identified by matching a list of experimental peptide masses (obtained generally by MALDI-MS) with the calculated list of all peptide masses of each entry in a database. Mass fingerprint method works well for isolated proteins, but the resulting protein identifications are not sufficiently specific for protein mixtures (e.g. for comigrating proteins).

The addition of sequencing capability to the MALDI method should make protein identifications more specific than those obtained by simple peptide-mass mapping. In stage 4, the peptide mixtures are separated by one or more steps of high-pressure liquid chromatography. In stage 5, a mass spectrum of the peptides eluting at this time point is taken (MS spectrum). The program generates a list of these peptides for fragmentation and a series of tandem mass spectrometric experiments (stage 6). MS-MS data consist in the isolation of a given peptide ion, fragmentation by energetic collision with gas, recording of the tandem or MS-MS spectrum and storing for matching against protein sequence databases. The outcome of the experiment is the identity of the peptides and therefore the proteins making up the purified protein population. Often MS-MS instruments are classified in one of two categories: tandem in space or tandem in time. Tandem in space instruments require a distinct analyser for each stage (isolation and fragmentation) of MS-MS. Today, almost all tandem in space MS-MS instruments are either triple quadrupole (QqQ) or hybrid instruments quadrupole/time-of-flight (Q/TOF). Trapping instruments are typically tandem in time. The various stages of MS-MS are performed in the same analyzer but separated in time.

Peptide Sequencing

In MS-MS experiments the isolated ions (termed parent ions) are induced to undergo a reaction that increases the internal energy of the ions, leading to dissociation. The ions resulting from the various reactions (product ions) are analyzed in the second stage of MS-MS. The scheme of a tandem MS experiment can be summarized in reaction in which m_p^+ is the parent ion, m_d^+ is the product or fragment ion and m_n the neutral fragment or another product ion if the parent ion is multiply charged:

$$m_p^+ \rightarrow m_d^+ + m_n$$

A crucial aspect of the MS-MS experiment is the reaction that occurs between the two MS stages. By far the most frequent reaction is unimolecolar dissociation, which is generally enhanced by some form of ion activation. The ion activation is necessary to increase the internal energy of the parent ion so that it will dissociate before analysis by MS^2. In practice, the activation cannot be separated from the dissociation, so the ion activation techniques are typically referred to as dissociation methods. The dissociation method almost universally used is collision-induced dissociation (CID) (24). In CID, the parent ion collides with a neutral target (collision) gas and some of the kinetic energy of the parent ion can be converted to internal energy. High energy CID spectra (keV) tend to be dominated by fragment ions formed by charge-remote mechanism (25), while low energy CID spectra (< 100 eV) often show intense fragment ions formed by abundant neutral losses. The CID spectra of peptides recorded at low collision energy tend to show abundant fragment ions formed by cleavage of the peptide bonds, while high-energy CID spectra often contain fragment ions formed by other backbone and also side-chain cleavages (26).

In mass spectrometric sequencing, the information that describes the amino acid sequence of a peptide is contained in a product ion spectrum. This product ion spectrum is obtained in a tandem mass spectrometry experiment by using collision induced dissociation of a protonated or multiply protonated peptide ion. Understanding the structure of protonated peptides and their fragmentation pathways plays key roles in one's ability to interpret ion spectra.

Peptide sequence identification by mass spectrometry involves fragmentation of a peptide to produce smaller m/z fragments; ideally, measured m/z values of these pieces can be assembled to produce the original sequence. Cleavage is commonly accepted to occur predominantly through charge-directed pathways. The mobile proton model (29) is a general description of the need to transfer a proton or protons intramolecularly to cleavage sites throughout the peptide. Considering, for example, a doubly protonated tryptic peptide, one proton will be localized on C-terminal Arg or Lys side chain and the second may be localized at one of the amide bonds or the N-terminal. In a collision process this second proton can be mobilized, giving a heterogeneous population of ions in which the second proton may reside at any one of the peptide bonds (mobile proton model). These transfers are facilitated by the proton affinity of heteroatoms. This interaction of the charge and heteroatoms is more important in the gas phase because the solvent molecules that typically stabilize the molecules in the solution phase are absent in the gas phase so internal groups in the molecule serve the role of "solvent" (28). Protonation of an amide nitrogen will weaken the amide bond and lead to cleavage, generating fragment ions. As result of the heterogeneous population of the precursor ion, a series of peptides bonds are cleaved, giving complementary series. A nomenclature (26) exists that is used to describe the fragment ion types that are produced by cleavage of different bonds along the peptide backbone and/or side chain. Cleavage of the backbone typically occurs at the peptide bond to produce **b** ions, if the amino terminal fragment retains the charge, or **y** ions, if the carboxy-terminal fragment retains the charge (fig. 2). In the case of multiply charged ions, a charge separation can occur to produce complementary ion pairs. Both partners of the complementary pair are not always detected in equal abundance, because they are not equally stable against further fragmentation or because instrument discrimination may enhance or diminish one partner of the pair. Although **b** and **y** ions are considered to be the most useful sequence ion types, because they correspond to cleavage of the amide bond, other ion types are observed and used in spectral interpretation. These include **a** ions which correspond formally to loss of CO from **b** ion; a Δm= 28 Da between two peaks suggests an **a-b** ion pair and is useful in identifying the ion series to which the peaks belong. The **y** series is sometimes accompanied by satellite peaks formally corresponding to NH_3 loss from the **y** ions, allowing designation of **y** ion series. Ions that correspond to immonium ions, or fragments of immonium ions, of individual amino acid residues in a peptide are often detected, even for residues from the internal portion of the sequence (29).

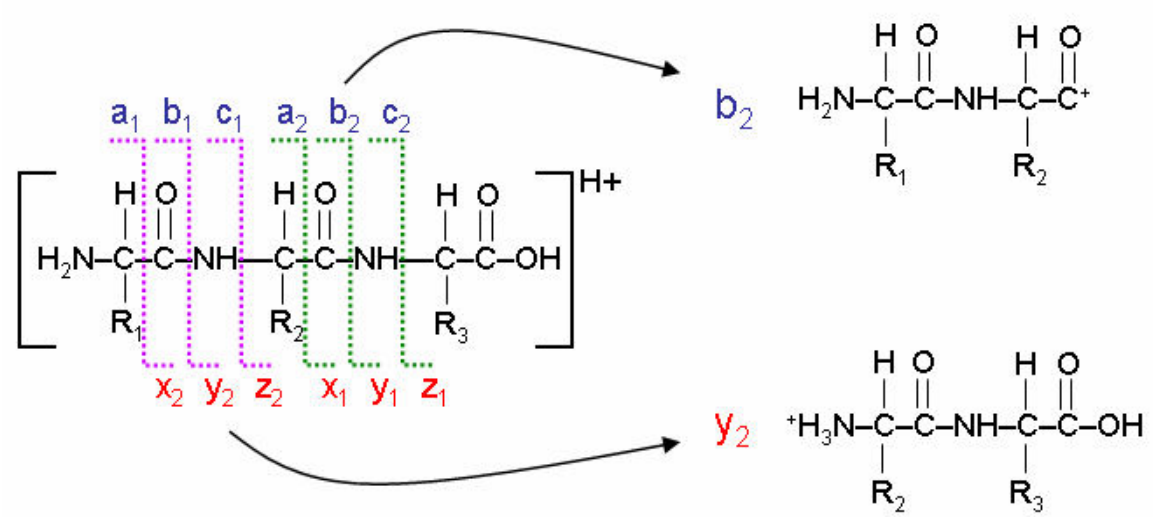


Figure 2. Schematic representation of nomenclature for fragmentation of peptide ions.

I.2.3 The 2nd Generation Technology

While 2D-GE still can be considered the most widely used separation technique prior to MS analysis, liquid chromatographic and, to a lesser extent, capillary electrophoresis separations are increasingly used in so-called "gel-free approaches". For this set-up, all proteins present in the sample are usually hydrolysed to obtain a very complex mixture that is then subjected to further separation steps.

It is worth considering that a single-dimension peptide chromatography does not provide sufficient peak capacity to separate peptide mixtures as complex as those generated by the proteolysis of protein mixtures of, for example, total cell lysates. To provide more peak capacity, various combinations of peptide separation schemes have been exploited. Most popular at present are two-dimensional (strong cation exchange/reversed phase) (30-31-32) or three-dimensional (strong cation exchange/affinity/reversed phase) (33) chromatographic separations of peptide mixtures generated by tryptic digestion of protein samples.

This strategy, takes advantage of the higher separation efficiency of chromatographic techniques at the peptide rather than the protein level.

In both MALDI- and ESI-MS, the relationship between the amount of analyte present and measured signal intensity is complex and partially understood. Mass spectrometers are therefore inherently poor quantitative devices.

Various and different approach for quantitative protein profiling have been proposed. The proven technique of stable-isotope dilution makes use of pairs of chemically identical analytes having different stable-isotope composition that can be differentiated in a mass spectrometer owing to their mass difference. The ratio of intensities of the peptide peaks in a given mass spectrum give a relative ratio of abundance of the two species. The labelling protocols can be broadly classified into two groups based on the method of incorporation of the quantitative tag: biological incorporation, where labelling of the peptide/protein is achieved by growing cells in media enriched in stable isotope-containing anabolites, and chemical incorporation, which relies on the use of a derivatization reagent for chemical modification of proteins in a site-specific manner after harvest of the proteins. The biological incorporation made use of ^{15}N-labeling in bacteria and yeast. Furthermore, the use of

the serine protease catalyzed attack of ^{18}O water in the digestion step originally used for the *de novo* sequencing of peptides (68-69) has also been applied to quantitative proteomics (70-71). More recently, Ong and coworkers (33) have developed a uniquely powerful method using stable isotope-containing amino acids in mammalian cell culture (SILAC). The use of ^{13}C substituted amino acids like ^{13}Carginine with SILAC affords some additional advantages for quantitative analyses with liquid chromatography mass spectrometry.

Chemical incorporation methods, called tagging (or labelling) strategies, targeted towards specific amino acid residues (cysteine, histidine and methionine) or N- or C-terminal peptides. This enables the enrichment of subfractions via affinity clean-up, resulting in the identification of an ever increasing number of proteins (35). A typical approach based on the isotope stable affinity tagging and MS is isotope coded affinity tag (ICAT). The reagent consist of a cysteine reactive group, a linker that contains either a heavy or light isotope and a biotin affinity tag (Fig.3).

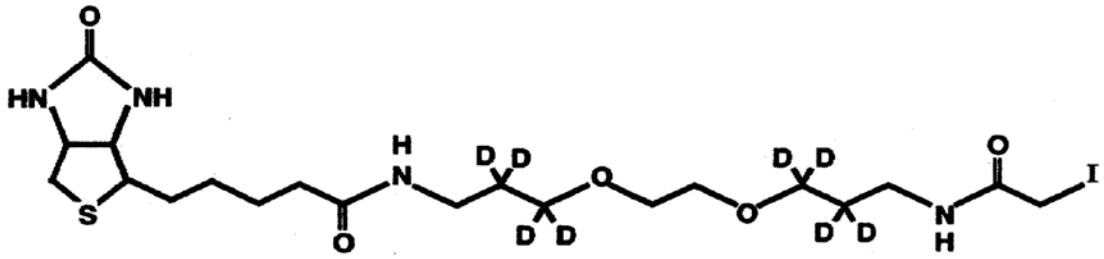


Figure 3. ICAT structure.

The affinity tag is attached, by chemical modification, to cysteine, allowing the sample to be purified by affinity chromatography (in this case, biotin-avidin chromatography). If a relatively rare amino acid, like, e.g. cysteine, is chosen as a target, only a relatively small fraction of peptides will be targeted, resulting in a significant reduction of sample complexity due to the affinity separation concept. The light and heavy reagents, chemically identical but mass differentiate stable isotope tags, are used to label the cysteine residues from proteins of two different sources. The labelled proteins are then combined and digested with trypsin. Cysteine-containing peptides are isolated using avidin affinity chromatography and subsequently identified, and quantified by MS. Each from one sample represent a quantitative standard for a chemically identical analyte from the other sample. The ratio of protein between the two states can then be determined by the measured peak ratio between the derivatized sample by two different form of ICAT reagent.

iTRAQ reagent technology is a newly developed method by Applied Biosystems for relative quantification of proteins that place isobaric mass labels at the N termini and lysine side chains of peptides in a digest mixture. The reagents are differentially isotopically labelled such that all derivatized peptides are isobaric and chromatographically indistinguishable, but yield signature or reporter ions following MS^2 that can be used to identify and quantify individual members of the multiplex set. The complete molecule consists of a reporter group (based on N-methylpiperazine), a mass balance group (carbonyl), and a peptide-reactive group (NHS ester). The overall mass of reporter and balance components of the molecule are kept constant using differential isotopic enrichment with ^{13}C, ^{15}N, and ^{18}O atoms (Fig.4). The reporter group ranges in mass from m/z 114.1 to 117.1, while the balance group ranges in mass from 28 to 31 Da, such that the combined mass remains constant (145.1 Da) for each of the four reagents. The labelled peptide are identical in mass and, therefore, also identical in single MS mode. In MS-MS mode low mass reporter

ion signals allow quantitation, while peptide fragment ion signals allow protein identification (36).

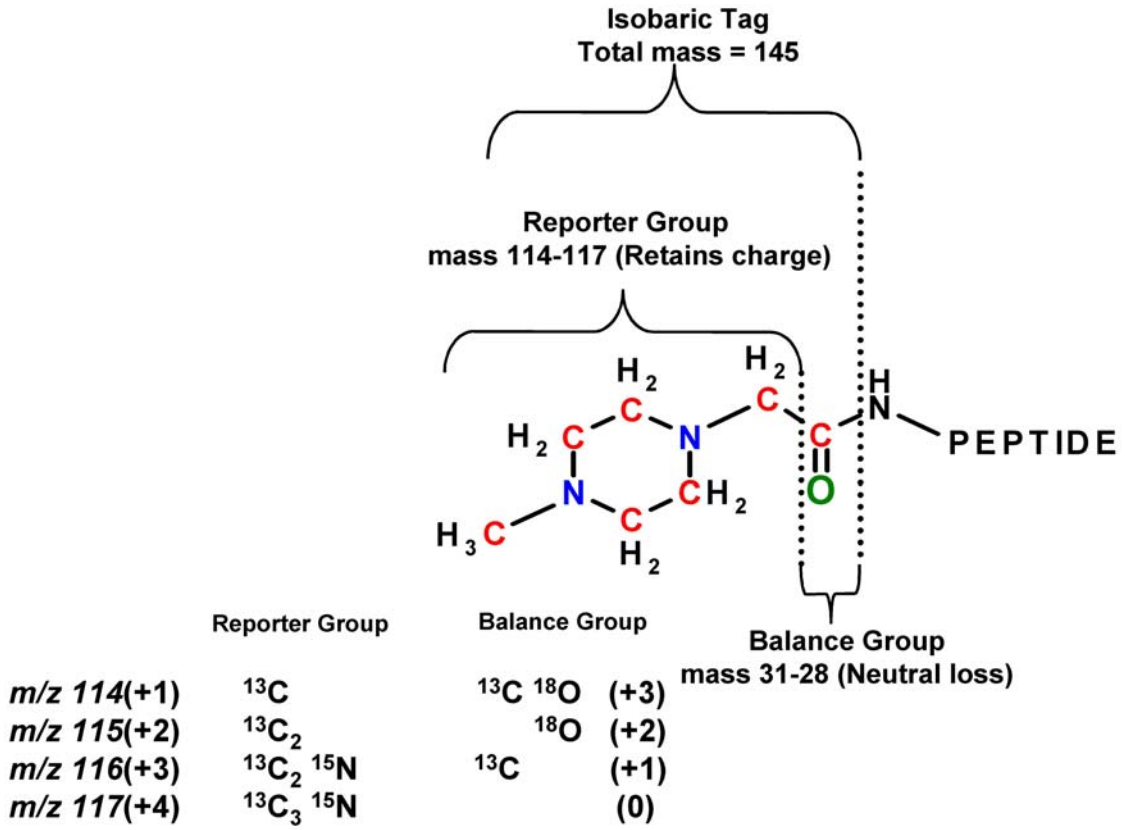


Figure 4. iTRAQ structure.

I.3 Functional proteomics: the crucial role of the phosphorylation

One of the descriptors of a protein, which is amenable to proteomics technology, is the delineation of post-translational modifications (PTMs). The delineation of a protein's function solely from a change in its abundance provides a limited view since numerous vital activities of proteins are modulated by PMTs that may not be reflected by changes in a protein abundance.

Protein phosphorylation is one of the most important and abundant PTM used to modulate protein activity and propagate signals within cellular pathways and networks (39). A primary role of phosphorylation is to act as a switch, to turn "on" or "off" a protein activity or a cellular pathway in reversible manner. This modification adjusts folding and function of proteins, e.g., enzymatic activities or substrate specificities, and regulating protein localization, complex formation and degradation. Cellular processes ranging from cell cycle progression, differentiation, development, peptide hormone response, metabolic maintenance and adaptation are all regulated by protein phosphorylation. The variety of functions in which phosphoproteins are involved necessitates a huge diversity of phosphorylations. Several amino acids can be phosphorylated by the four known types of phosphorylation. Serine, threonine and

tyrosine residues but also unusual amino acids such as hydroxy-proline can be O-phosphorylated. N-, S- and acyl- phosphorylation are far less spread and occur mostly on histidine and lysine (N-phosphates), cysteine (S-phosphates) and aspartic and glutamic acid residues (acyl-phosphates). The O-phosphorylations are stable under acidic conditions, whereas N-phosphorylations are stable under alkaline conditions. Acyl-phosphates are very reactive molecules and are labile in acid as well as in alkali. Phosphocysteine shows a moderate stability under acidic and alkali conditions (40).

Moreover, protein phosphorylation is generally present at substoichiometric level in the cell (41). In fact, whereas some residues are always quantitatively phosphorylated, others may only be transiently modified (42).

Protein phosphorylation is tightly regulated by a complex set of kinases and phosphatases the enzyme responsible for protein phosphorylation and dephosphorylayion (43-44). It has been estimated that as many as one–third of the proteins within a given eukaryotic proteome undergoes reversible phosphorylation and there are more than 100000 estimated phosphorylation sites in the human proteome (45). The cooperation of kinases and phosphatases is highly dynamic, intensely regulated and phosphorylation cycle may take place on a very shot timescale. In addition to the problems concerning regulation, the analysis is complicated by the complexity of phosphorylation patterns: in many cases, the effects of phosphorylation are combinatorial and multiple sites are phosphorylated .

Thus, the highlighting of phosphoproteome is a huge and challenging task with regard to the dynamics and different kinds of phosphorylation generating a variety of phosphoproteins that are not accessible to a single analytical method.

I.4 Phosphoproteome Tools

I.4.1. Initial efforts

Although two-dimensional polyacrylamide gel electrophoresis (2D-GE) has been used primarily for the separation and quantification of the relative abundance of proteins from different systems, methods are also being developed to characterize the phosphoproteome. One classic approach for characterizing protein phosphorylation relies on (^{32}P)-labeling, followed by two-dimensional gel electrophoresis (46-47). Typically gel spots of interest containing phosphorylated proteins are excised, digested, and analyzed by Edman sequencing or mass spectrometry.

A major advantage of this technique relies in the detection of both N-,S- and O- phosphorylation. The signals can be absolutely quantified. Introduction of radioactive phosphate-group can be done *in vivo* or *in vitro.* The latter is carried out by incubating, for an appropriate time, the protein with a chosen kinase and (γ-^{32}P)-ATP. Major drawbacks of *in vitro* labelling techniques relies in side reactions leading to some unspecific phosphorylations due to very high reagent concentrations and reaction conditions. Introduction of radioactive phosphate group *in vivo* is achieved by application of (γ-^{32}P)-ATP to the tissue, cell ect. However, artificial phosphorylations can occur because of the cellular stress induced by the incorporated radioactivity for instance causing DNA damage and thus activating DNA

repair system or promoting apoptosis. Furthermore, constitutively phosphorylated proteins with low phosphate turn over rate incorporate only very small amount of radioactivity and may thereby escape detection.

The quantitative problem is approached by comparing difference in relative abundance of phosphoproteins, ^{32}P-labeled, visualized by autoradiography measuring relative intensities of the spot on the gel (48). Nevertheless, the ^{32}P labelling is not the method of choice for high-throughput proteome-wide analysis because of issues with handling radioactive compounds and the associated contamination of analytical instrumentation.

As an alternative approach, commercially available phospho-stains are used. They are less sensitive than radioactive methods but the handling of these “inactive” reagents is more convenient so far. As an example, Pro-Q Diamond, (Molecular Probes) a fluorescent dye, allows to detect 1-20 ng of phosphoprotein in polyacrylamide gel, depending on the phosphorylation state of the protein. For individual phosphoproteins, the strength of the signal correlates with the number of phosphate groups. Depending on the staining principle, this method is mostly specific for Ser/Thr- phosphorylations.

Antibodies can be used to discriminate among serine, threonine and tyrosine phosphorylation. The specifity and sensitivity of immunostaining are strongly dependent on the respective antibodies. Various phosphotyrosine antibodies of good specifity are available and only little cross-reactivity is observed (50). To measure changes in protein phosphorylation state, two distinct proteome sample can be run and immunoblotted and the intensity of spot can be compared.

Over the years, several strategies have been developed that increase the sensitivity of phosphoproteome analysis and eliminate the need for radioactive and antibody labelling. MS-based methods have been developed that provide more effective tools to identify, and potentially quantify, specific sites of phosphorylation.

I.4.2. Mass spectrometry and phosphorylation on the rise

Recently, mass spectrometry based methods have emerged as powerful and preferred tools for the analysis of post-translational modifications including phosphorylation due to higher sensitivity, selectively, and speed than most biochemical technique (53-54).

Once a protein has been isolated, several technique can be used to identify and localize the modified amino acids. In some case, the precise measurement by mass spectrometry of the molecular weight of the intact protein, can address the average number of the modified residues just by looking at the increment of the protein mass compared with the unmodified one. However, the exact location and nature of the modifying group cannot be achieved just by mass spectrometric measurement of the intact protein. As an example a shift of 80-Da can be interpreted as arising by the addition of a phosphate group or may be due to Gly→His exchange generated by a point mutation.

To better characterize the phosphorylated residues of the protein it is necessary to analyze the peptide generated by enzymatic (usually trypsin) or chemical degradation of the protein.

Phosphopeptides may be identified simply by examination of the list of observed peptide masses for mass increases of 80 Da compared with the list of

expected peptide masses. Any ambiguities can be resolved by sequencing the peptide using tandem mass spectrometry (MS/MS) (52-53).

Although this method is relatively straightforward, it also misses many phosphorylated peptides because (51) peptide maps are frequently incomplete, even for non-phosphorylated proteins (some tryptic peptides are poorly ionized or poorly recovered (53)), the increased acidity of the phosphate group generally results in decreased ionization efficient of peptide (54), and competition for ionization peptides in a mixture results in suppression of signal for some peptide (54). To confirm, the identity of a phosphopeptide, a simple phosphatase reaction will cause a downward shift in mass of multiples 80 Da for each phosphopeptide. Furthermore, removing phosphate group could enhance the relative intensity of these peptides (67).

A mass spectrometry based approach to phosphopeptide analysis (55-56) is especially powerful when used in conjunction with electrospray ionization (ES) (63). Combined with on-line liquid chromatography, ESI-MS has formed the basis for several novel techniques for identifying phosphopeptides. In this last case, the ion suppression phenomena is reduced by separation of the peptides prior to ionization. This method has added benefit of concentrating dilute samples and removing salt that interferes with the ionization process.

Phosphopeptide detection by peptide mapping preceding and following enzymatic dephosphorylation circumvents the poor ionization efficiency of phosphopeptides. Larsen et al. have reported enzymatic dephosphorylation of complex peptide mixtures directly on a matrix assisted laser desorption ionization (MALDI) target (58).

Moreover, phosphopeptides have characteristic fragmentation patterns. In fact, PSD is a process where specific ions, called metastable, decompose in the flight tube because they are not sufficiently stable. In the case of serine and threonine phosphorylated peptides, the most common fragmentations are due to the loss of H_3PO_4, more abundant, and HPO_3. These fragments don't appear in the spectrum at the exact masses because they are not properly focused by mirror.

However, despite the advances in mass spectrometry for the phosphorylation analysis, some difficulties still remain. First, the generally low phosphorylation stoichiometry of most of the proteins such that phosphopeptides are essentially present in low amount in the generated complex peptide mixtures. Second, the increased hydrophilicity and hence reduced retention of phosphopeptides on reversed-phase materials. Finally, the selective suppression of their ionization/detection efficiencies in the presence of large amounts of unphosphorylated peptides during mass spectrometry analysis in positive mode.

I.4.3. Enrichment of phosphoproteins and phosphopeptides

Separation technologies such as affinity, liquid reverse phase, ion exchange chromatographies and capillary electrophoresis prior to MS analysis have been gradual used in proteomics to enrich phosphoproteins that may otherwise be lost in detection. The enrichment step combined with the high sensitivity of MS technologies provides great potential for phosphoproteome characterization.

Enrichment at level of phosphoproteins and/or phosphopeptide becomes increasingly important when dealing with complex mixture.

Specific antibodies can be used to enrich phosphoproteins by immunoprecipitation (66) from complex lysates. In this case, the extraction of the

phosphoproteins led to the substantial simplification of the protein pattern and the enrichment of the low abundant phosphoproteins.

Recently, the use of miniaturized immobilized metal affinity chromatography (IMAC), in which phosphopeptides are bound noncovalently to resins that chelate Fe (III) or other trivalent metals, followed by base elution, has proved to be a potentially valuable method in phosphopeptide enrichment. 21–31. With further refinement, this technique may offer the best performance for large-scale phosphorylation analysis. (59).

Further several methods for selective enrichment of phosphoproteins and phosphopeptides use chemical modification of the phosphate group (64). This approach does not distinguish between O-glycosylated and phosphorylated analytes, therefore, requiring additional experiments to confirm phosphorylation. Zhou *et al.* (65) established a multi-step derivatization method that is capable of enriching not only Ser/Thr-phosphorylated but all types of phosphorylated peptides. Phospho-peptides are bound to a sulfidryl-containing compound via phosphoamidate-bonds and can thereby be covalently linked to a solid support with immobilized iodocetyl-groups. The phosphate-groups are not cleaved off the respective residues, so native phosphopeptides are obtained after elution with TFA. A drawback of this method is the expensive work that has to be done to block all amino- and carboxy-groups within the peptides quantitatively to prevent intramolecular and intermolecular condensation.

I.5 Aim of the PhD thesis

This PhD thesis focuses on the development of new proteomic approaches based on classical biochemical methods coupled to new mass spectrometry methods to study the protein phosphorylation. Phosphoproteome studies aimed at comprehensive analysis of protein phosphorylation are based on a multi-task protocol addressing the identification of the phosphoproteins, exact localization of the phosphorylation sites, and preferably quantification of the extent of the phosphorylation.

Due to low stoichiometry, heterogeneity, and low abundance, enrichment of phosphopeptides is an important step of this analysis. Thus, a part of this thesis is focused on the development of new enrichment methods coupled to mass spectrometry.

A final part of the work aimed at the development of a strategy to study phosphorylation without preliminary enrichment by exploiting the high performances of a novel hybrid mass spectrometer, equipped with a Linear Ion Trap. To prove their wide enforceability these methodologies were first optimized by using a model phosphoprotein, bovine α-casein, and then applied to different biological systems.

I.6 References

1. O'Donovan C, Apweiler R, Bairoch A. The human proteomics initiative (HPI). Trends Biotechnol. 2001. 19(5):178-81.

2. Corthals GL, Wasinger VC, Hochstrasser DF, Sanchez JC. The dynamic range of protein expression: a challenge for proteomic. Electrophoresis. 2000. 21(6):1104-15.
3. Washburn MP, Wolters D, Yates JR 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol. 2001. 19(3):242-7.
4. Harry JL, Wilkins MR, Herbert BR, Packer NH, Gooley AA, Williams KL. Proteomics: capacity versus utility. Electrophoresis. 2000. 21(6):1071-81.
5. Gygi SP, Corthals GL, Zhang Y, Rochon Y, Aebersold R. Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. Proc Natl Acad Sci U S A. 2000. 15;97(17):9390-5.
6. Garfin DE. Two-dimensional gel electrophoresis: an overview. Trends in Analytical Chemistry. 2003. Vol. 22: No 5.
7. O'Farell PH, High resolution two-dimentional electrophoresis of proteins. J Bio Chem. 1975. CCL: 4007-4021.
8. Mann M, Hendrickson RC, Pandey A. Analysis of proteins and proteomes by mass spectrometry. Annu Rev Biochem. 2001. 70:437-73.
9. Berggren KN, Chernokalskaya E, Lopez MF, Beechem JM, Patton WF. Comparison of three different fluorescent visualization strategies for detecting Escherichia coli ATP synthase subunits after sodium dodecyl sulfate-polyacrylamide gel electrophoresis Proteomics. 2001. 1(1):54-65.
10. Rabilloud T. Two-dimensional gel electrophoresis in proteomics: old, old fashioned, but it still climbs up the mountains. Proteomics. 2002. 2(1):3-10.
11. Unlu M, Morgan ME, Minden JS. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts Electrophoresis. 1997. 18(11):2071-7.
12. Gauss C, Kalkum M, Lowe M, Lehrach H, Klose J. Analysis of the mouse proteome. (I) Brain proteins: separation by two-dimensional electrophoresis and identification by mass spectrometry and genetic variation. Electrophoresis. 1999. 20(3):575-600.
13. Pandey A. & Mann M. Proteomics to study genes and genomes. Nature. 2000. 405, 837–846.
14. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Electrospray ionization for mass spectrometry of large biomolecules. Science. 1989. 6;246(4926):64-71.
15. Pandey A. & Mann M. Proteomics to study genes and genomes. Nature. 2000. 405, 837–846.
16. Aebersold R, Goodlett DR. Mass spectrometry in proteomics. Chem Rev. 2001. 101(2):269-95.
17. Henzel WJ, Billeci TM, Stults JT, Wong SC, Grimley C, Watanabe C. Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. Proc Natl Acad Sci U S A. 1993. 1;90(11):5011-5.
18. Pappin DJ, Hojrup P, Bleasby AJ. Rapid identification of proteins by peptide-mass fingerprinting. Curr Biol. 1993. 1;3(6):327-32.
19. Yates JR 3rd, Speicher S, Griffin PR, Hunkapiller T. Peptide mass maps: a highly informative approach to protein identification. Anal Biochem. 1993. 1;214(2):397-408.
20. Clauser KR, Hall SC, Smith DM, Webb JW, Andrews LE, Tran HM, Epstein LB, Burlingame AL. Rapid mass spectrometric peptide sequencing and mass

matching for characterization of human melanoma proteins isolated by two-dimensional PAGE. Proc Natl Acad Sci U S A. 1995. 23;92(11):5072-6.
21. Gooley AA, Ou K, Russell J, Wilkins MR, Sanchez JC, Hochstrasser DF, Williams KL. A role for Edman degradation in proteome studies. Electrophoresis. 1997. 18(7):1068-72.
22. Mann M, Wilm M. Error-tolerant identification of peptides in sequence databases by peptide sequence tags. Anal Chem. 1994. 15;66(24):4390-9.
23. Mann M. A shortcut to interesting human genes: peptide sequence tags, expressed-sequence tags and computers. Trends Biochem Sci. 1996. 21(12):494-5.
24. McLuckey SA, Goeringer DE, Glish GL. Collisional activation with random noise in ion trap mass spectrometry. Anal Chem. 1992. 1;64(13):1455-60.
25. Glish GL and Vachet RW. The basics of mass spectrometry in the twenty-first century. Nat Rev Drug Discov. 2003. 2(2):140-50.
26. Johnson RS, Martin SA, Biemann K, Stults JT, Watson JT. Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. Anal Chem. 1987. 1;59(21):2621-5.
27. Herrmann KA, Somogyi A, Wysocki VH, Drahos L, Vekey K. Combination of sustained off-resonance irradiation and on-resonance excitation in FT-ICR. Anal Chem. 2005. 1;77(23):7626-38.
28. Wysocki VH, Resing KA, Zhang Q, Cheng G. Mass spectrometry of peptides and proteins. Methods. 2005. 35(3):211-22.
29. Takach EJ, Hnes WM, Patterson DH, Juhasz P, Falick AM, Vestal ML, Martin SA.Accurate mass measurements using MALDI-TOF with delayed extraction. J Protein Chem. 1997. 16(5):363-9.
30. Washburn MP, Wolters D, Yates JR 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol. 2001. 19(3):242-7.
31. Link AJ, Eng J, Schieltz DM, Carmack E, Mize GJ, Morris DR, Garvik BM, Yates JR 3rd. Direct analysis of protein complexes using mass spectrometry. Nat Biotechnol. 1999. 17(7):676-82.
32. Han DK, Eng J, Zhou H, Aebersold R. Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. Nat Biotechnol. 2001. 19(10):946-51.
33. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics. 2002. 1(5):376-86.
34. Wang H, Hanash S. Multi-dimensional liquid phase based separations in proteomics. J Chromatogr B Analyt Technol Biomed Life Sci. 2003. 5;787(1):11-8.
35. Tao WA, Aebersold R. Advances in quantitative proteomics via stable isotope tagging and mass spectrometry. Curr Opin Biotechnol. 2003. 14(1):110-8.
36. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet-Jones M, He F, Jacobson A, Pappin DJ. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics. 2004. 3(12):1154-69.
37. Wilkins MR, Pasquali C, Appel RD, Ou K, Golaz O, Sanchez JC, Yan JX, Gooley AA, Hughes G, Humphery-Smith I, Williams KL, Hochstrasser DF. From

proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. Biotechnology (N Y). 1996. 14(1):61-5.
38. Aebersold R, Goodlett DR. Mass spectrometry in proteomics. Chem Rev. 2001. 101(2):269-95.
39. Ben-Levy R, Leighton IA, Doza YN, Attwood P, Morrice N, Marshall CJ, Cohen P. Identification of novel phosphorylation sites required for activation of MAPKAP kinase-2. EMBO J. 1995.1;14(23):5920-30.
40. Meyer HE, Eisermann B, Heber M, Hoffmann-Posorske E, Korte H, Weigt C, Wegner A, Hutton T, Donella-Deana A, Perich JW. Strategies for nonradioactive methods in the localization of phosphorylated amino acids in proteins. FASEB J. 1993. 7(9):776-82.
41. Hunter T. When is a lipid kinase not a lipid kinase? When it is a protein kinase. Cell. 1995. 6;83(1):1-4.
42. Schlessinger J. Cellular signaling by receptor tyrosine kinases. Harvey Lect. 1993-1994;89:105-23.
43. Venter JC, et. al. The sequence of the human genome. Science. 2001 16;291(5507):1304-51. Erratum in: Science 2001. 5;292(5523):1838.
44. Lander ES, et. al. Initial sequencing and analysis of the human genome. Nature. 2001. 15;409(6822):860-921. Erratum in: Nature 2001. 2;412(6846):565. Nature 2001. 7;411.
45. Kalume DE, Molina H, Pandey A. Tackling the phosphoproteome: tools and strategies. Curr Opin Chem Biol. 2003. 7(1):64-9.
46. Kaufmann H, Bailey JE, Fussenegger M. Use of antibodies for detection of phosphorylated proteins separated by two-dimensional gel electrophoresis. Proteomics. 2001. 1(2):194-9.
47. Bendt AK, Burkovski A, Schaffer S, Bott M, Farwick M, Hermann T. Towards a phosphoproteome map of Corynebacterium glutamicum. Proteomics. 2003. 3(8):1637-46.
48. Cohen P. The regulation of protein function by multisite phosphorylation--a 25 year update. Trends Biochem Sci. 2000. 25(12):596-601.
49. Berwick DC, Tavare JM. Identifying protein kinase substrates: hunting for the organ-grinder's monkeys. Trends Biochem Sci. 2004. 29(5):227-32.
50. McLachlin DT, Chait BT. Analysis of phosphorylated proteins and peptides by mass spectrometry. Curr Opin Chem Biol. 2001. 5(5):591-602.
51. Annan RS, Huddleston MJ, Verma R, Deshaies RJ, Carr SA. A multidimensional electrospray MS-based approach to phosphopeptide mapping. Anal Chem. 2001. 1;73(3):393-404.
52. Biemann K. Sequencing of peptides by tandem mass spectrometry and high-energy collision-induced dissociation. Methods Enzymol. 1990. 193:455-79.
53. Pandey A, Andersen JS and Mann M. Use of Mass Spectrometry to Study Signaling Pathways. Sci STKE. 2000. 2000(37):PL1.
54. Janek K, Wenschuh H, Bienert M, Krause E. Phosphopeptide analysis by positive and negative ion matrix-assisted laser desorption/ionization mass spectrometry. Rapid Commun Mass Spectrom. 2001. 15(17):1593-9.
55. Poulter L, Ang SG, Gibson BW, Williams DH, Holmes CF, Caudwell FB, Pitcher J, Cohen P. Analysis of the in vivo phosphorylation state of rabbit skeletal muscle glycogen synthase by fast-atom-bombardment mass spectrometry. Eur J Biochem. 1988. 15;175(3):497-510.
56. Ohguro H, Palczewski K, Ericsson LH, Walsh KA, Johnson RS. Sequential phosphorylation of rhodopsin at multiple sites. Biochemistry. 1993. 1;32(21):5718-24.

57. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Electrospray ionization for mass spectrometry of large biomolecules. Science. 1989. 6;246(4926):64-71.
58. .Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal Chem. 1988. 15;60(20):2299-301.
59. Gaberc-Porekar V, Menart V. R Perspectives of immobilized-metal affinity chromatography. J Biochem Biophys Methods. 2001. 30;49(1-3):335-60 B)
60. Andersson L, Porath J. Isolation of phosphoproteins by immobilized metal (Fe3+) affinity chromatography. Anal Biochem. 1986. 154(1):250-4.
61. Olcott MC, Bradley ML, Haley BE. Photoaffinity labeling of creatine kinase with 2-azido- and 8-azidoadenosine triphosphate: identification of two peptides from the ATP-binding domain. Biochemistry. 1994. 4;33(39):11935-41.
62. Ficarro SB, McCleland ML, Stukenberg PT, Burke DJ, Ross MM, Shabanowitz J, Hunt DF, White FM. Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. Nat Biotechnol. 2002. 20(3):301-5.
63. Ballif BA, Villen J, Beausoleil SA, Schwartz D, Gygi SP. Phosphoproteomic analysis of the developing mouse brain. Mol Cell Proteomics. 2004. 3(11):1093-101.
64. F Oda Y, Nagasu T, Chait BT. Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. Nat Biotechnol. 2001. 19(4):379-82.
65. Zhou H, Watts JD, Aebersold R. A systematic approach to the analysis of protein phosphorylation. Nat Biotechnol. 2001. 19(4):375-8.
66. Marcus K, Immler D, Sternberger J, Meyer HE. Identification of platelet proteins separated by two-dimensional gel electrophoresis and analyzed by matrix assisted laser desorption ionization-time of flight-mass spectrometry and detection of tyrosine-phosphorylated proteins. Electrophoresis. 2000. 21(13):2622-36.
67. Petrilli P, Pucci P, Morris HR, Addeo F. Assignement of phosphorylation sites in buffalo beta casein by fast atom bombardment mass spectrometry. Biochem Biophys Res Commun. 1986. 140(1):28-37
68. Schnolzer M, Jedrzejewski P, Lehmann WD. Protease-catalyzed incorporation of 18O into peptide fragments and its application for protein sequencing by electrospray and matrix-assisted laser desorption/ionization mass spectrometry. Electrophoresis. 1996. 17:945–953.
69. Shevchenko A, Chernushevich I, Ens W, Standing KG, Thomson B, Wilm M, Mann M. Rapid 'de novo' peptide sequencing by a combination of nanoelectrospray, isotopic labeling and a quadrupole/time-of-flight mass spectrometer. Rapid Commun. Mass Spectrom. 1997. 11:1015–1024.
70. Mirgorodskaya OA, Kozmin YP, Titov MI, Korner R, Sonksen CP, Roepstorff P. Quantitation of peptides and proteins by matrix-assisted laser desorption/ionization mass spectrometry using (18)O-labeled internal standards. Rapid Commun. Mass Spectrom. 2000. 14(14):1226-32.
71. Yao X, Freas A, Ramirez J, Demirev PA, Fenselau C. Proteolytic 18O labeling for comparative proteomics: model studies with two serotypes of adenovirus Anal. Chem. 73(13):2836-42. Erratum in: Anal Chem. 2004. 1;76(9):2675.

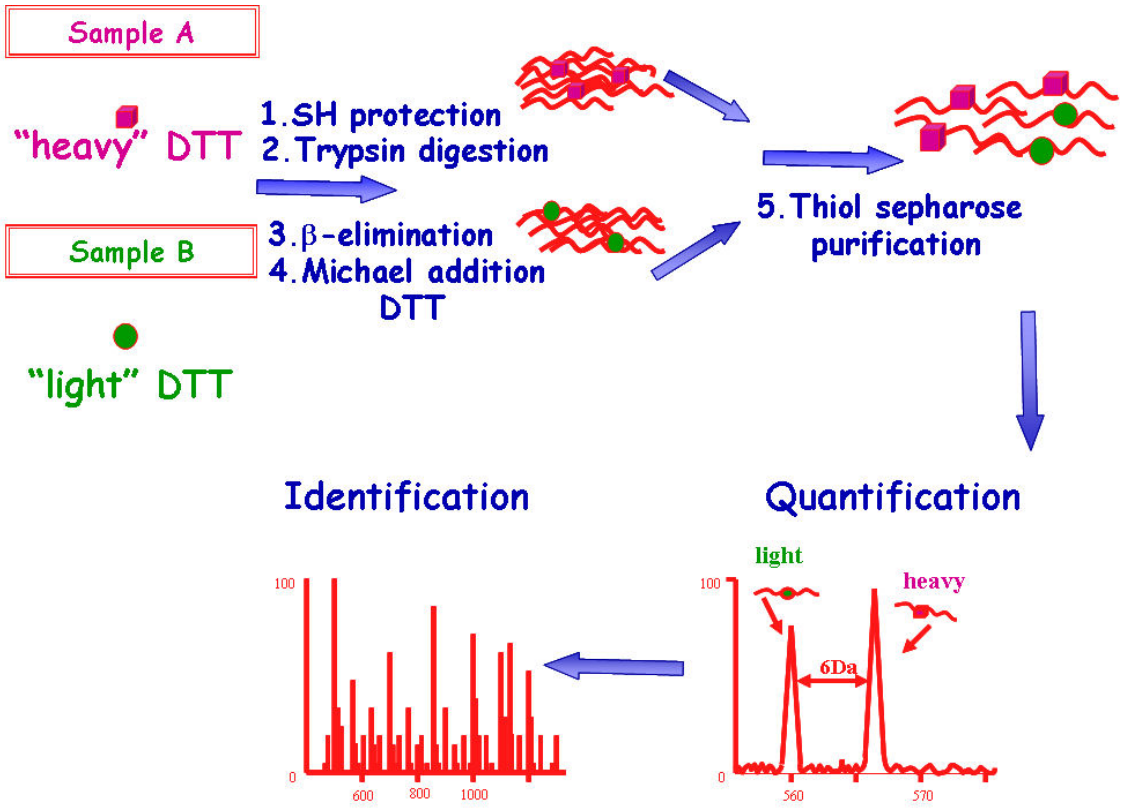
II. Mapping phosphorylation sites: a new strategy based on the use of isotopically labelled DTT and mass spectrometry.

II.1 Introduction

The main methods that have been used in order to identify phosphorylation sites by MS involve the chemical modification of phosphopeptides. Usually the modification requires a β-elimination reaction in an alkaline environment to remove the phosphate group from phosphoserine and phosphothreonine, followed by a Michael addition reaction with different nucleophiles. The method was first introduced by Meyer to sequence phosphopeptides by Edman degradation (1) and since then many different nucleophilic agents have been used such as ethanethiol (2-3), ethanedithiol (4-5). A different derivatization method suitable for the analysis of tyrosine, serine and threonine phosphopeptides requiring several steps has been described by Zhou (6).

New methods that involve modifying phosphoproteins with affinity tags in combination with stable isotope incorporation (7-8) have been developed for the specific enrichment and quantitation of phosphopeptides (see chapter 1).
Gygi and co-workers have recently reported (9) an approach for quantitative differential analysis that employs isotope coded affinity tags (ICAT).

Here, an improved approach, based on simple chemical manipulations and mass spectrometric procedures (see Scheme 1), for the selective analysis of phosphoserine and phosphothreonine in protein mixtures following conversion of the peptide phosphate moiety into DTT derivatives is described. The new residues in the sequence were shown to be stable and easily identifiable under general conditions for tandem mass spectrometric sequencing applicable to the fine localization of the exact site of phosphorylation. The suitability of the method with the principle of stable isotope coding is also demonstrated. However the major aim of the present work is devoted to the use of isotopically labelled DTT thus allowing a simple and direct quantitative MS analysis.



Scheme1. Strategy to map phosphorylation sites by labelling DTT

II.2 Materials and Methods

TPCK-treated Trypsin, dithiothreitol, barium hydroxide and alfa-cyano-4-hydroxycinnamic acid were from Sigma-Aldrich (St. Louis, MO, USA). Deuterated dithiothreitol was purchased from CEA (Saclay, France). Sepharose thiopyridil resin was from Pierce Biotechnology (Rockford, IL, USA). All other reagents and solvents were of the highest purity available from Carlo Erba (Milan, Italy).

In situ digestion

Proteins were fractionated by SDS-electrophoresis on 12.5% polyacrylamide gels under reducing conditions

The analysis was performed on the Comassie blue-stained proteins excised from gels. The excised spots were washed first with ACN and then with 0.1M ammonium bicarbonate. Protein samples were reduced by incubation in 10mM dithiothreitol (DTT) for 45 min at 56°C. The cysteines were alkylated by incubation in 5mM iodoacetamide for 15 min at room temperature in the dark. The gel particles were then washed with ammonium bicarbonate and ACN.

Enzymatic digestion was carried out with trypsin (12.5ng/µl) in 50mM ammonium bicarbonate pH 8.5 at 4°C for 4 hours. The buffer solution was then removed and a new aliquot of the enzyme/buffer solution was added for 18 hours at 37°C. A minimum reaction volume, enough for the complete rehydratation of the gel was used. Peptides were then extracted washing the gel particles with 20mM

ammonium bicarbonate and 0.1% TFA in 50% ACN at room temperature and then lyophilised.

Phosphate group modification

β-elimination reactions were carried out by incubating the peptide mixture in 0.1 M $Ba(OH)_2$, at 37°C, for 90 min under nitrogen. Carbonic dioxide was then added. The precipitated barium carbonate was then removed by centrifugation at 13000 g/min for 5 min. Then DTT, in the light and heavy form, 30% w/v in Hepes buffer 10mM pH 7.5 was added for 3 hr, at 50°C under nitrogen.

Isolation and enrichment of tagged peptides

After addition reaction the peptide mixtures were treated with activated thiol sepharose resin directly in batch. The resin was washed in water twice and then with binding buffer, 0.1M Tris-HCl, pH7.5. The peptide mixture was added to the resin and washed with the binding buffer. The modified peptides were then extracted by using the elution buffer, 20mM DTT in 10mM Tris-HCl, and pH 7.5.

MS analyses

MALDI mass spectra were recorded using an Applied Biosystem Voyager DE-PRO Instrument operating in reflector mode. A mixture of analyte solution and α-cyano-hydroxycinnamic acid (10mg/ml in 70% acetonitrile, 0.1% trifluoroacetic acid) was applied to the metallic sample plate and dried at room temperature. Mass calibration was performed using external peptide standards by Applied Biosistems. Raw data were analysed using the computer software provided by the manufacturer and reported as monoisotopic masses. MSMS spectra were recorded using an Applied Biosystem 4700 Proteomics Analyzer mass spectrometer operating in reflector mode.

Peptide mixture were analysed by LCMS on a single quadrupole ZQ electrospray (Waters Micromass) coupled to an HPLC (2690 Alliance purchased by Waters) on a Phenomenex 30 mm x 0.46 mm i.d. reverse-phase. Peptide mixture were eluted at a flow rate of 0.2 ml/min. with a 5-65% acetonitrile-water gradient in 60 min. Data were acquired and processed by centroiding an isotopic distribution using the Mass-Lynx program (Micromass).

II.3 Results and Discussion

This study reports a novel approach to phosphoprotein mapping based on site-specific modification of phosphoseryl/phosphotheonyl residues. The validity of its applicability for the identification of the phosphorylation sites in peptides was tested by MALDI TOF/TOF mass spectrometer. Here, it is shown that an alternative conversion of the P-Ser and P-Thr residues into compounds, which are stable during collision induced dissociation, provided easily interpretable product ion spectra.

To test this procedure under realistic but controlled conditions, it is carried out the reactions by using phosphorylated bovine α-casein, a normal protein utilised in these studies (10-11-9). The protein samples after separation from SDS polyalcrylamide gel were reduced, alkylated and digested *in situ* with trypsin (see

Material and Methods section). The peptide mixtures were then directly analysed by MALDI-MS (Table1). The phosphopeptides were identified by their 80 Da mass difference compared to the peptides expected from the sequence for the presence of the phosphate moiety. As an example, the signal at m/z 1660.6 was assigned to the peptide 106-119 within α-casein sequence carrying a phosphate group as reported in figure 1. The peptide mixture was submitted to the strategy described above (see scheme 1) and each step was monitored via MALDI-MS. Reaction conditions were optimised for both the beta elimination and DTT addition reactions.

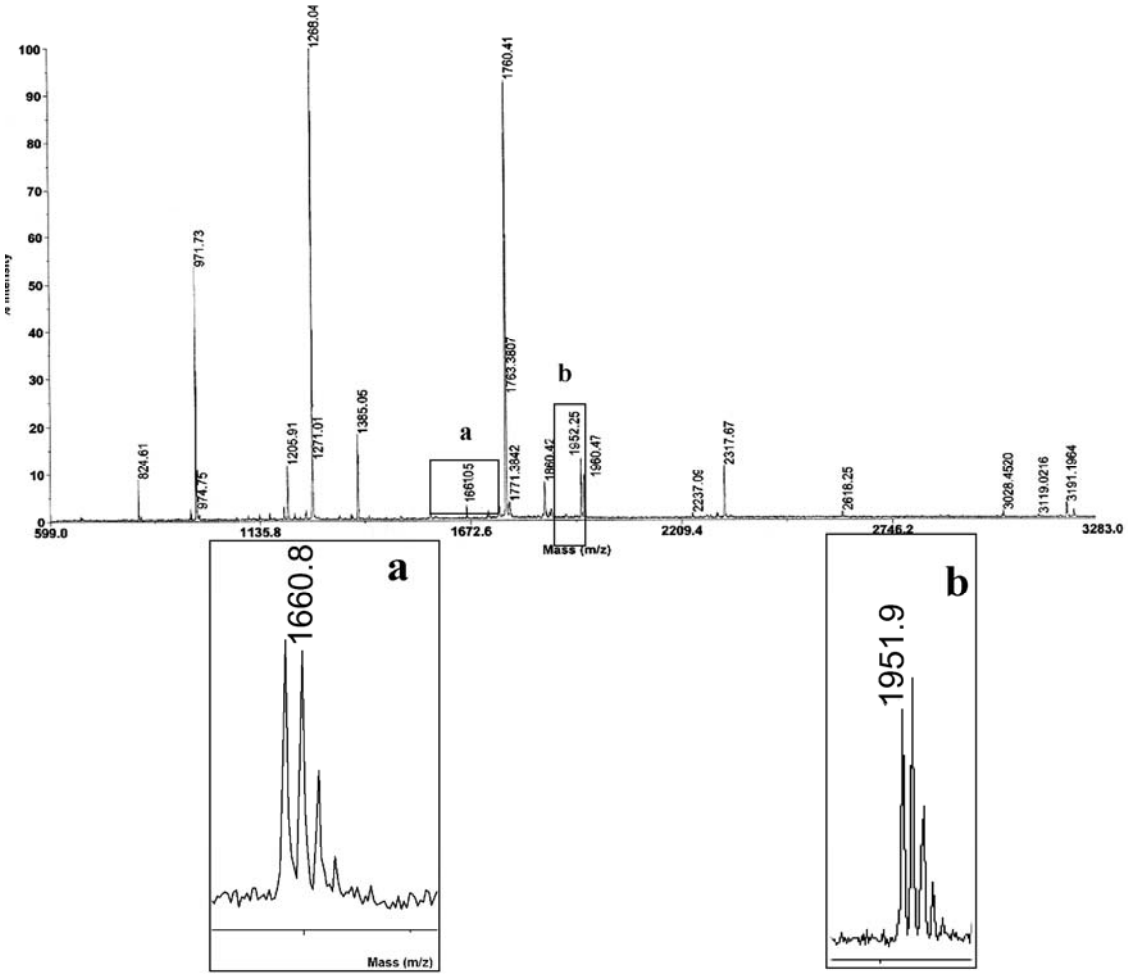


Figure 1. MALDI-MS analysis of the trypsin peptide mixture from α-casein. The insets report the amplification of the region of the spectrum showing the phosphorylated peptides.

Table 1. MALDI-MS analysis of the peptide mixture of bovine α-casein digested with trypsin.

Experimental MH^+	Theoretical MW	Peptide	Note
3026.3	3025.4	166-193	
2316.31	2315.13	133-151	
1952.3	1950.94	104-119	1P
1759.7	1758.93	8-22	
1660.5	1659.8	106-119	1P
1384.5	1383.73	23-34	
1337.7	1336.6	80-90	
1267.5	1266.69	91-100	
946.3	945.5	35-42	
910.5	909.46	125-132	
875.6	874.55	1-7	
831.3	830.37	84-90	
748.5	747.3	194-199	
615.38	614.32	120-124	

II.3.1 β-elimination reaction

Under strong alkaline conditions the phosphate moiety on Ser-P and Thr-P undergoes to β-elimination to form dehydroalanine (ΔSer) or dehydroalanine-2-butyric acid (ΔThr) respectively. In the case of O-glycosylated peptides, the corresponding glycopeptides have to be separated by a simple lectin affinity chromathography step. The α,β-unsatured residues are potent Michael acceptor, which can readily react with a nucleophile (12-13).

In the present study, the phosphate moieties were removed via barium hydroxide ion-mediated β-elimination from pSer and pThr. The barium hydroxide was used in place of the previously reported NaOH (14-15) taking in account the higher purity of this reagent and the higher reactivity towards the pThr residues. To get read of the removal reagent excess we simply made use of solid carbon dioxide. The use of this reagent permitted to remove the barium carbonate as a pellet on the bottom of a vial after centrifugation. It is north noting that when the pellet of barium carbonate was dissolved and analysed by MALDI-MS, no peptide was detected in the mass spectrum thus indicating the absence of a sort of aspecific precipitation. On the contrary, the mass spectral analysis carried out on the supernatant showed the occurrence of a series of signals corresponding to the β-eliminated peptides. As an example, the signal at m/z 1854.3 was assigned to the peptide 104-119 within α-casein occurring 98 Da lower than that expected on the basis of the amino acid sequence, thus indicating that the pSer115 was converted in dehydroalanyl after β-elimination reaction (Fig. 2).

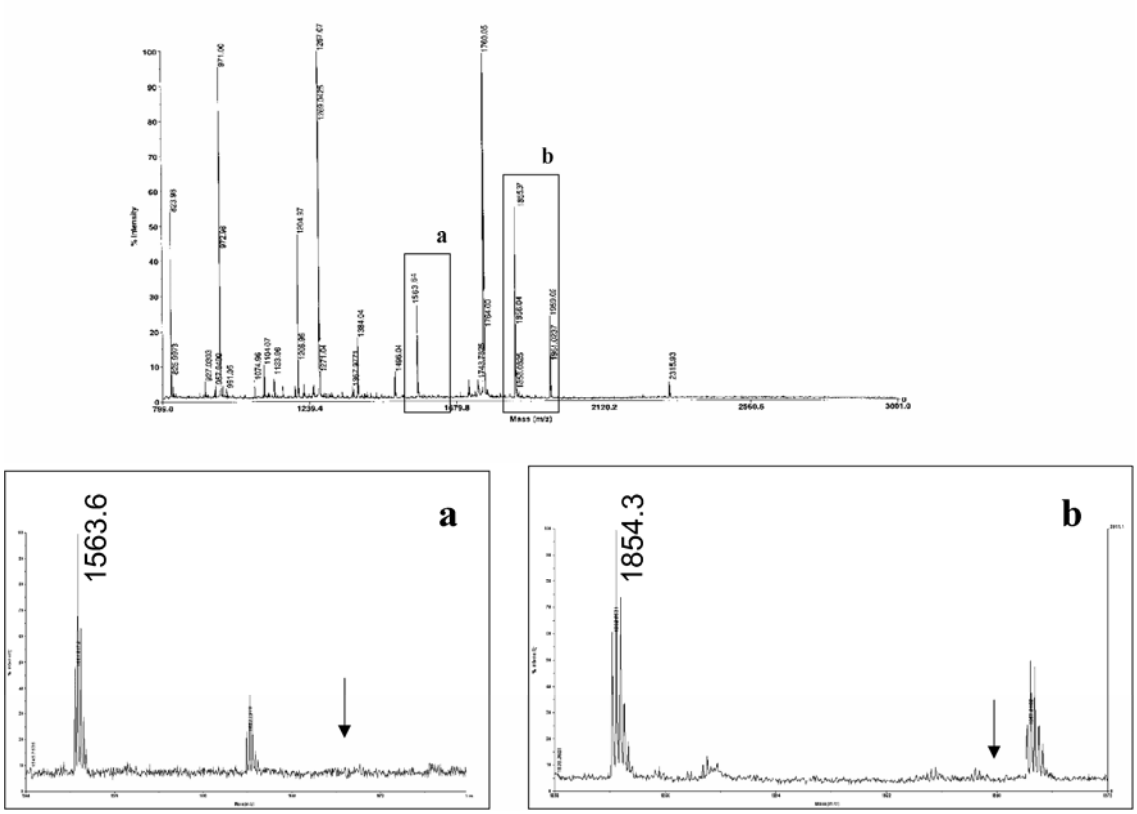


Figure 2. MALDI-MS analysis of trypsin peptide mixture from α-casein following β-elimination reaction. The boxes report amplifications of the region of the spectrum showing the β-eliminated peptides. The arrows indicate the disappearance of the phosphorylated peptides and the occurrence of the new β-eliminated signals occurring at minus 98Da.

II.3.2. DTT as a bifunctional reagent

The β-eliminated peptide mixture was submitted to a Michael-type addition by using dithiothreitol (DTT) by following the procedure described (12). The rational behind the choice of this reagent is due to its higher solubility and reactivity with respect to the already proposed EDT (13-15). In fact EDT has limited solubility in aqueous medium, thereby requiring sample cleanup prior LCMS analysis since the conversion is carried out in the presence of substantial amount of organic solvent (16). The addition reaction results in the creation of a free thiol group in place of what was formerly a phosphate moiety. The extent of reaction was monitored via MALDIMS thus indicating that yield of the addition reaction thus revealing that the amount of DTT-modified peptide was about 70 % (Fig 3), as expected for a typical Michael reaction addition (12). A comparison made using EDT showed that even in the presence of a very large reagent excess yields were not higher than 30% (data not shown). However, the direct mass spectral analysis of a complete peptide mixture suffered of suppression phenomena leading to the unappearance of signals due to most of the phosphorylated peptides.

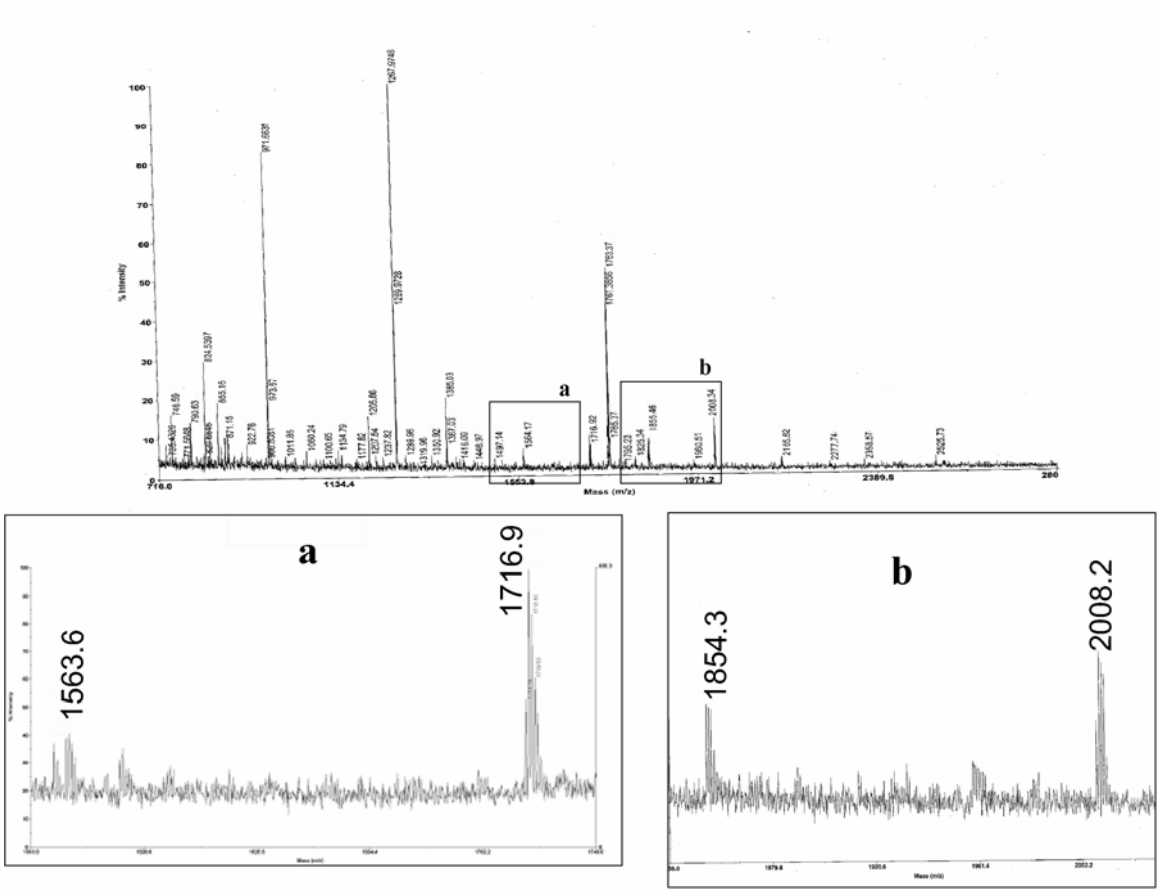


Figure 3. MALDI-MS analysis of trypsin peptide mixture of the protein after DTT labelling reaction. Insets: enlarged region of the spectrum showing the β-eliminated and DTT labelled peptides.

II.3.3. Affinity capture of thiolated peptides

Specific enrichment of thiolated peptides has easily achieved by using a sepharose thiopyridil resin. MALDI-MS analysis performed on the peptides eluted from the affinity purification step reveals now major mass signals at m/z 1716.9, 2008.3, 2039.1 and 3001.2. These values were assigned to the β-eliminated fragments 106-119, 104-119, 43-58 and 59-79 modified by 1, 1, 2 and 5 DTT adducts respectively as indicated in Figure 4.

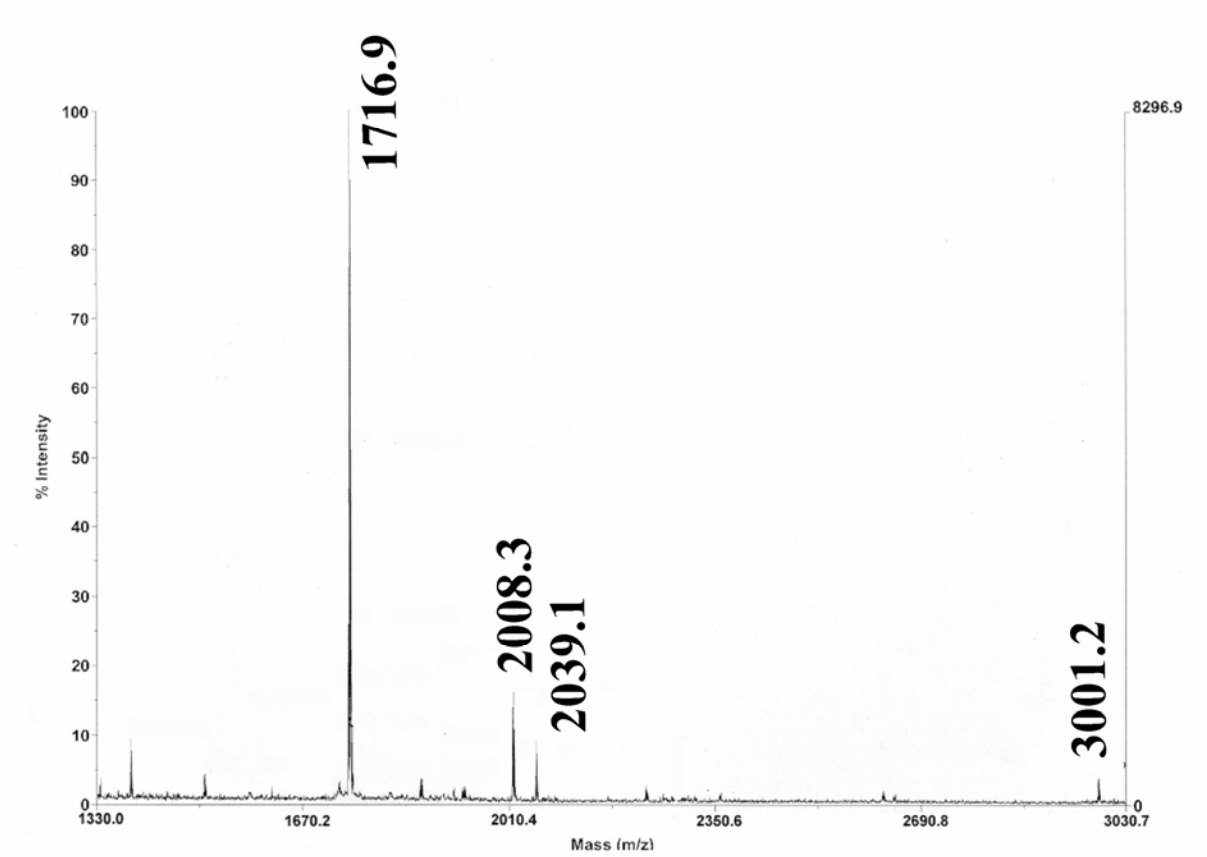


Figure 4. MALDI-MS analysis of DTT labelled peptides following purification step. Each signal is attributed to the corresponding modified peptide according to Table 2.

It is worth nothing that the described procedure resulted in the identification of all the phosphorylated tryptic peptides, including the pentaphosphorylated one that has escaped in different attempts used (9-10-11). The data are summarised in Table 2.

Table 2. MALDI-MS analysis of the purified DTT labelled peptides of the triptic digest of the α-casein.

Experimental MH^+	Theoretical MW	Peptide	Note
3001.2	2999.9	59-79	5 DTT
2039.1	2038.68	43-58	2 DTT
2008.3	2006.94	104-119	1 DTT
1716.9	1715.8	106-119	1 DTT

II.3.4. MS/MS analysis

Sequence information can be obtained either by MALDI-TOF in the PSD mode or by MALDI TOF/TOF. The fine localization of the phosphorylation sites in the peptides was achieved by MS/MS sequencing. The conversion of P-Ser and P-Thr in DTT tagged residues originate compounds that are stable under high-energy collisions in the MALDI TOF/TOF instrument. In fact, the tandem mass spectral

analyses by MALDITOF/TOF performed on the precursor ion at m/z 2008.5 (Figure 5) from the tryptic mixture of α-casein after labelling with DTT revealed the presence of diagnostic fragment ions.

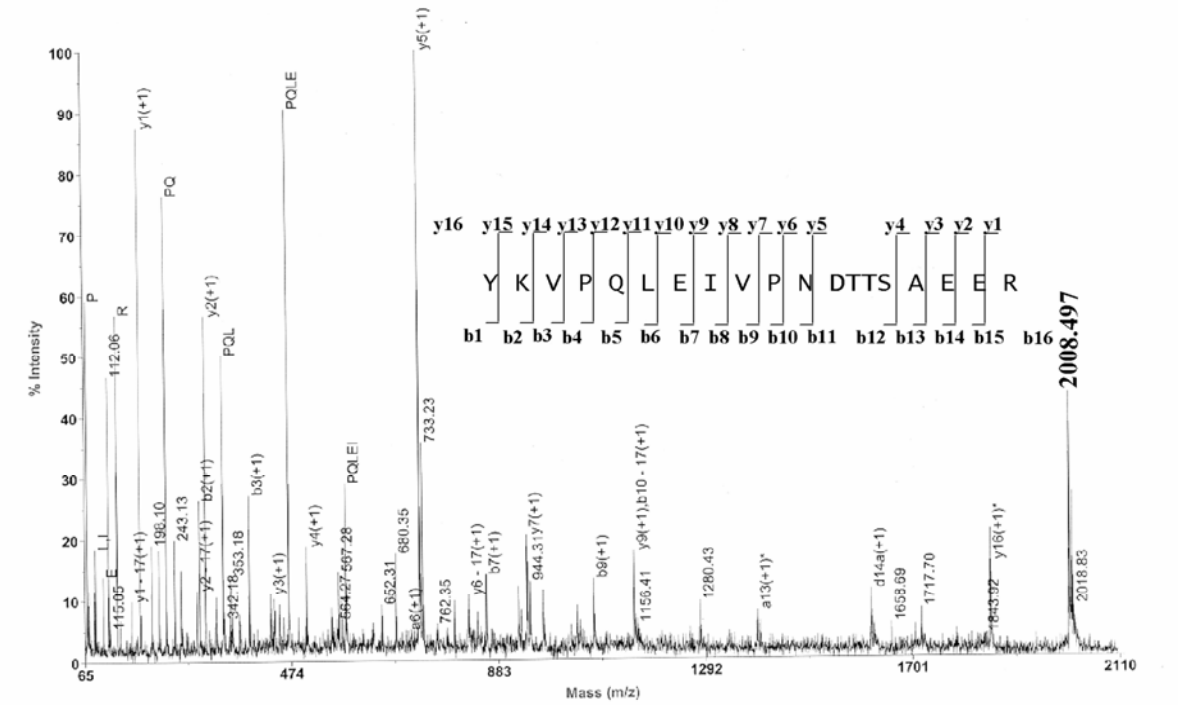


Figure 5. Tandem mass spectrometry identification of α-casein phosphopeptide. The Ser residue containing the DTT label is identified by the loss of m/z 223.4 between the y4 and y5 ions of the modified peptide.

The derivatized product ions were identified from either the corresponding b or y product ion series. The MS-MS spectrum revealed the occurrence of the complete series of N-terminal fragment ions b1-b15 plus the C-terminal y1-y16 for the modified phosphopeptide with DTT labelled phosphoserine residue. The fragment ions y4 and y5, in fact, show a mass difference of 223 Da corresponding to a β-eliminated serine residue, labelled with DTT, thus confirming the phosphorylation site at Ser115.

II.3.5 Differential isotope coded analysis

To address the problem of quantification (9-17-18), having created a thiol group at the phosphoserine/phosphothreonine site, it is possible to follow basically the strategy outlined by Gygi et al. (9). Because of the fact that we had the availability of fully deuterated DTT, we exploited a much simpler isotope tagging strategy. The peptides were quantified by measuring, in the mixture, the relative signal intensities for pairs of peptide ions of identical sequence differentially labelled with light or heavy DTT. The mass difference within the heavy DTT reagent (6 Da) reliably allows a good separation between the molecular ions. As an example the mass signals, recorded in the MALDI spectrum, at m/z 2008.5 and 2014.5 were attributed to the DTT labelled peptide 104-119 modified with the DTT light and heavy form respectively. Pairs of peptides tagged with the light and heavy DTT reagents are chemically identical and therefore serve as ideal mutual internal standards for quantitation. The ratio between the intensities of the lower and upper mass

components of these peaks provide an accurate measurement of the relative abundance of each peptide and hence of the related proteins in the original cell pool because the MS relative intensity of a given peptide is independent of the isotopic composition of the defined isotopically tagged reagent.

To further illustrate the ability of this labelling strategy in quantifying the relative phosphorylation of a peptide from two different samples, DTT and DTT-D_6 were used to label samples containing stoichiometric concentrations of α-casein in ratios of 1:1, 2:1 and 4:1. The peptide mixture were analysed by MALDIMS operating in reflector mode. The spectra were obtained by acquiring about 2000 shot/spectrum in order to produce a "stable" peak. In our experience this is the average number of shots to have an even distribution of shots on the entire well surface, thus allowing a confident averaging independently of the peptide content of the single crystals. As shown for the DTT peptide 104-109 (Figure 6), the ratios automatically integrated for each mass spectrum resulting in excellent agreement with the concentration used in the labelling experiment (Table 3).

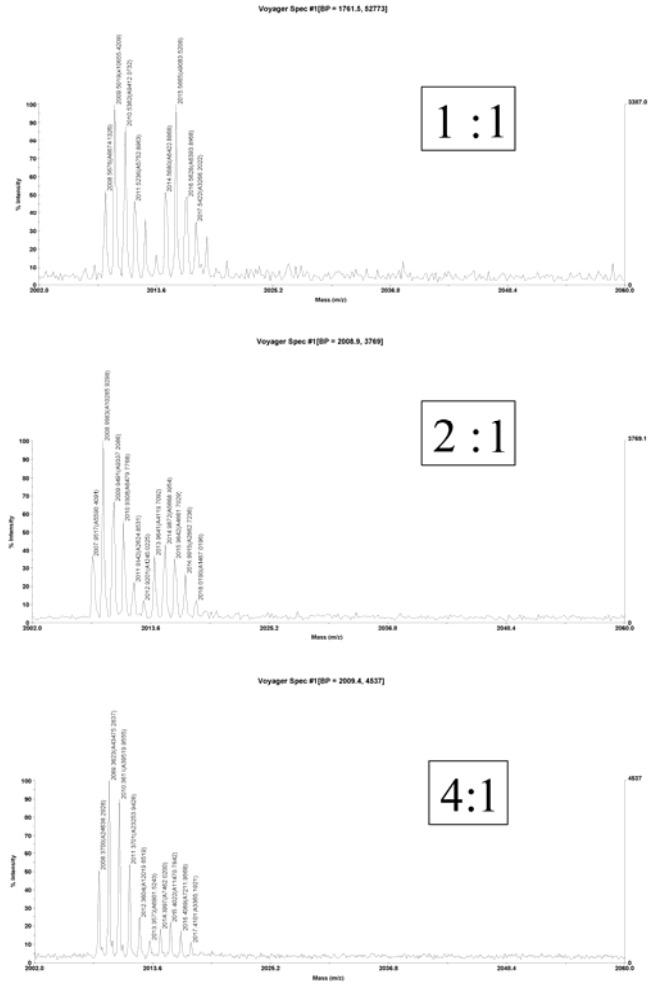
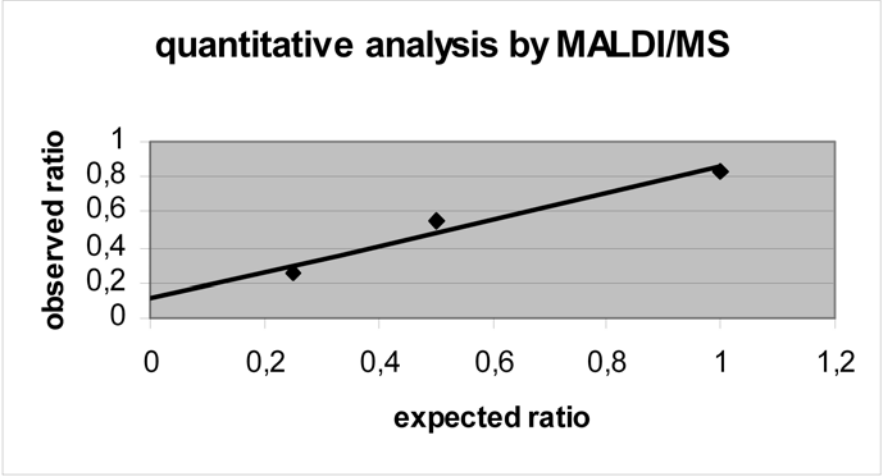


Figure 6. Stoichiometric conversion of the phosphorylated states using DTT-D_0/D_6 labelling via MALDI-MS. Partial MALDI-MS spectra of the modified peptide 104-119. Samples of α-casein containing ratios of 1:1, 2:1 and 4:1 were labelled with DTT-D_0 :DTT-D_6, combined, isolated via DTT strategy reported in the Scheme 1 and directly analysed via MALDI-MS.

Table 3. Quantitative analysis by MALDI-MS.

Expected ratio	DTT (D0) labelled peak area	DTT (D6) labelled peak area	Observed Ratio
1: 1	29168.1	24306.7	1.2: 1
2: 1	34318.05	18779.51	1.8: 1
4: 1	142907.11	36411.46	3.9: 1



Y=0.7429x + 0.11 $R^2 = 0.9565$

An aliquot of the same peptide mixture was submitted to a more rigorous quantitative analysis via LCMS (Figure 7) as reported in Table 4, demonstrating the correctness of the previous MALDI-MS data.

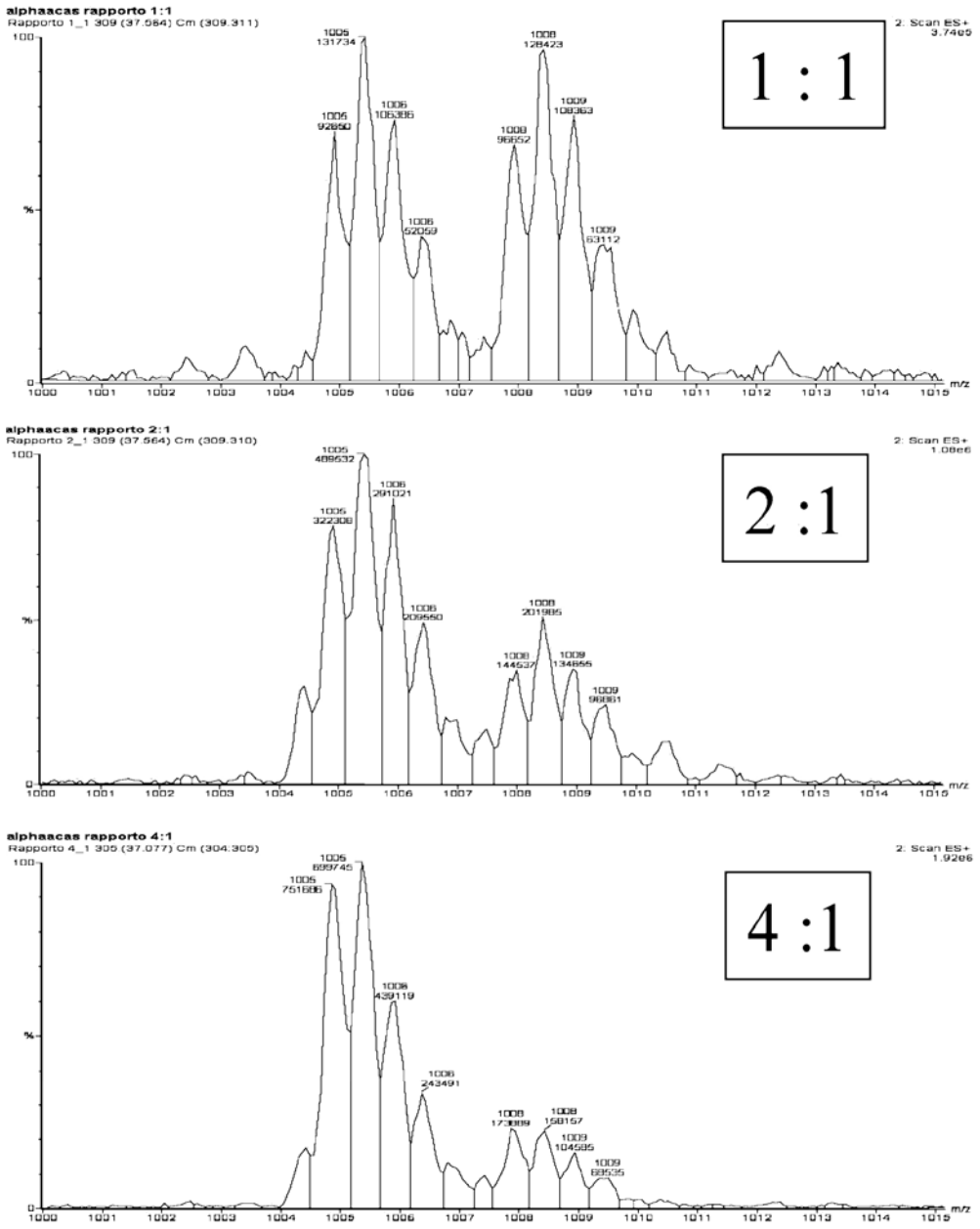
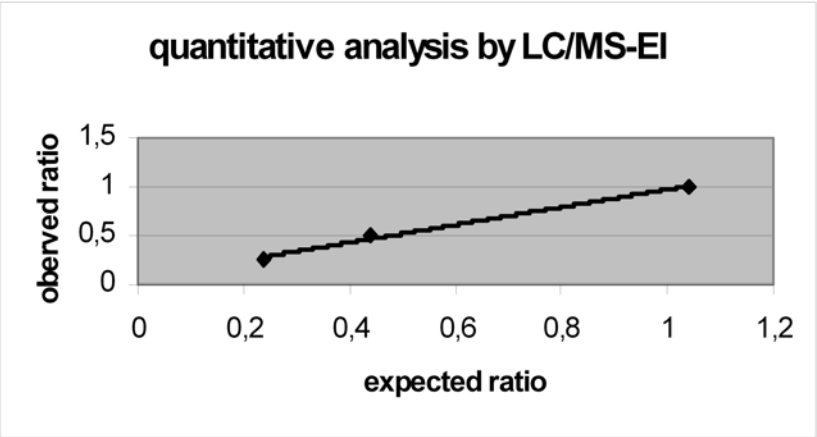


Figure 7. Stoichiometric conversion of the phosphorylated states using DTT-D_0/D_6 labelling via ESI-MS. LC-ESMS spectrum of the doubly charged ion of the modified peptide 104-119. Samples of α-casein containing ratios of 1:1, 2:1 and 4:1 were labelled with DTT-D_0 :DTT-D_6, combined, isolated via DTT strategy reported in the Scheme 1 and analysed via LC-ESIMS.

Table.4 Quantitative analysis by LC-ESI-MS.

Expected ratio	DTT (D_0) labelled peak area	DTT (D_6) labelled peak area	Observed ratio
1: 1	382829	396550	0.96: 1
2: 1	1312411	578038	2.27: 1
4: 1	2134041	505166	4.22: 1



$Y=0.9135x + 0.0596$ $R^2 = 0.9918$

II.4 Conclusions

This study is an integrated simple methodology, which involves chemical replacement of the phosphate moieties by affinity tags. This method to locate and quantitate phosphorylated residues may result of interest in signalling pathways and control mechanisms studies involving phosphorylation or dephosphorylation of serine/threonine residues. It is worth considering that such a quantitative approach can be extended to the labelling of the O-glycosylation sites. In fact, the O-glycopeptides retained on a lectin affinity column can be eluted and submitted to the procedure described by Wells and coworkers (19). The quantitative analysis can be performed by using light and heavy forms of DTT to selective label the modified sites thus leading to a quantitative description of O-glycosylation related to different cell growth conditions.

II.5.References

1. Meyer HE, Hoffmann-Posorske E, Korte H, Heilmeyer LJ. Sequence analysis of phosphoserine-containing peptides. Modification for picomolar sensitivity. FEBS Lett. 1986. 204:61–66.
2. Resing KA, Johnson RS, Walsh KA. Mass spectrometric analysis of 21 phosphorylation sites in the internal repeat of rat profilaggrin, precursor of an intermediate filament associated protein. Biochemistry. 1995. 34:9477–9487.
3. Jaffe H, Veeranna T, Pant HC. Characterization of serine and threonine phosphorylation sites in beta-elimination/ethanethiol addition-modified proteins by electrospray tandem mass spectrometry and database searching. Biochemistry. 1998. 37:16211–16224.
4. Goshe MB, Conrads TP, Panisko EA, Angell NH. Phosphoprotein isotope-coded affinity tag approach for isolating and quantitating phosphopeptides in proteome-wide analyses. Anal. Chem. 2001. 73:2578–2586.

5. Adamczyk M, Gebler JC, Wu J. Identification of phosphopeptides by chemical modification with an isotopic tag and ion trap mass spectrometry. Rapid Commun. Mass Spectrom. 2001. 15:1481–1488.
6. Zhou H, Watts JD, Aebersold R. A systematic approach to the analysis of protein phosphorylation. Nat. Biotechnol. 2001. 19:375–378.
7. Jarvik JW, Telmer CA. Epitope tagging. Annu. Rev. Genet. 1998. 32:601-18.
8. Smolka MB, Zhou H, Purkayastha S and Aebersold R. Optimization of the Isotope-Coded Affinity Tag-Labelling Procedure for Quantitative Proteome Analysis. Anal. Biochem. 2001. 297: 25-31.
9. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nature Biotechnol. 1999.17:994-9.
10. Goshe MB, Conrads PT, Panisko EA, Angell NH, Veenstra TD and Smith RD. Phosphoprotein Isotope-Coded Affinity Tag Approach for Isolating and Quantitating Phosphopeptides in Proteome-Wide Analyses. Anal. Chem. 2001. 73: 2578-86.
11. Goshe MB, Veenstra TD, Panisko EA, Conrads TP, Angell NH and Smith RD. Phosphoprotein Isotope-Coded Affinity Tags: Application to the Enrichment and Identification of Low-Abundance Phosphoproteins. Anal. Chem. 2002. 74: 607-16.
12. Goering HL, Relyea DI and Larsen DW. Preparation of thiols: formation from alkenes. J. Amer. Chem. Soc. 1956. 78: 348.
13. Conrads TP, Issaq HJ, Veenstra TD. New tools for quantitative phosphoproteome analysis. Biochem. Biophys. Res. Commun. 2002. 25: 885.
14. Li W, Boykins RA, Backlund PS, Wang G, Chen HC. Identification of phosphoserine and phosphothreonine as cysteic acid and beta-methylcysteic acid residues in peptides by tandem mass spectrometric sequencing. Anal. Chem. 2002. 15: 5701.
15. Mega T, Nakamura N, Ikenaka T. Modifications of substituted seryl and threonyl residues in phosphopeptides and a polysialoglycoprotein by beta-elimination and nucleophile additions. J. Biochem. 1990. 107: 68.
16. Jaffe H, Veeranna H, Pant C. Characterization of serine and threonine phosphorylation sites in beta-elimination/ethanethiol addition-modified proteins by electrospray tandem mass spectrometry and database searching. Biochemistry. 1998.17: 16211.
17. Mann M. Quantitative proteomics? Nat. Biotechnol. 1999. 17: 954.
18. Oda Y, Huang K, Cross FR, Cowburn D, Chait BT. Accurate quantisation of protein expression and sitespecific phosphorylation. Proc. Natl. Acad. Sci. USA. 1999. 96: 6591.
19. Wells L, Vosseller K, Cole RN, Cronshaw JM, Matunis MJ and Hart GW. Mapping sites of O-GlcNAc modifications using affinity tags for serine and threonine post-translational modifications. Mol. Cell. Proteomics. 2002. 1: 791.

III. Phosphoproteomics Applications by Improved Immobilized Metal Affinity Chromatography and Selective tandem mass spectrometry

III.1 History and principle of IMAC

In the last few years, a number of strategies devoted to the separation of phosphorylated proteins and peptides from the non-phosphorylated counterparts have been exploited. These include phosphoresidue enrichment strategies such as affinity purification by metal affinity chromatography (IMAC).

Firstly introduced by Porath and coworkers in 1975, IMAC was originally termed metal chelate affinity chromatography, for use as a group separation method (1). IMAC represents an affinity separation method based on the coordination between peptides/proteins and metal ions on a solid support (matrix).

Beginning in the 1990s, IMAC was increasingly employed as a pre-fractionation tool in the field of proteomic research (2).

To date, the most popular metal ions used in IMAC are Cu^{2+}, Ni^{2+}, Zn^{2+}, Co^{3+}, Fe^{3+} and Ga^{3+} (3). Transition ions bound resine-immobilize chelating ligands available from several suppliers. Commonly used chelating groups are bidentate (e. g. iminodiacetic acid, IDA, and aminohydroxamic acid), tridentate (e.g. dipicolyamine), tetradentate (carboxymethylated aspartic acid and nitrilotriacetic acid) or pentadentate (N,N,N'-tris(carboxymethyl)ethylendiamine) chelating ligands according to the number of occupied coordination bonds (4). Appropriate combination of chelating ligands and metal ions is crucial for selectivity and adsorption efficiency of IMAC in separating target fractions, because chelating ligands adversely affect protein retention in chromatography (3).

Chelating compounds are linked by means of spacers with solid supports, which hold physicochemical characteristics favourable for IMAC application: specific adsorption, good stability, high porosity, high flow rates and no degradation upon regeneration. Furthermore, these supports are easy to derive.

Traditional stationary phases were based on gel matrices, such as agarose or crosslinked dextran. Some amino acid residues that have electron-donating side chains, especially those containing surface-exposed atoms of N, S, O, have high affinities for a series of metal ions. Specifically, exposed histidines are the common anchors for the binding of proteins to the column (4). Other residues, such as glutamic acid, aspartic acid, and tyrosine, can contribute to the metal binding.

In addition to the types of chelators and properties of metal ions and proteins, other factors in mobile phase, such as pH, ionic strength and buffer composition, affect protein-metal binding too.

The binding of proteins to IMAC can be processed at a given pH in which the electron donor groups on the protein surface are unprotonated. The pH value used in the binding buffer of IMAC cover a wide range depending on the used metal.

The target proteins or peptides are eluted from IMAC by changing pH or using high concentrations of imidazole (competitive binding), based on the nature of the interaction.

III.2 Quadrupole scan modes and phosphoproteomics

Owing to its reliability, speed and sensitivity, tandem mass spectrometry has become a valuable tool for peptide sequence analysis and PTM characterization. A limitation of this approach for phosphopeptide analysis is the large dynamic range between the non phosphorylated and phosphorylated peptides. After IMAC enrichment, phosphopeptides are separated by nanoflow reversed-phase high-pressure liquid chromatography (RP-nHPLC) with online nano-electrospray ionization for subsequent tandem mass spectrometry (RP-nHPLC/ESI/MS-MS).

Phosphopeptides undergo characteristic fragmentation pathways when subjected to collision-induced dissociation (CID) allowing them to be distinguished from non-phosphorylated peptides. In negative ion mode, for example, phosphopeptides fragment producing marker ions at m/z 79 (PO_3^-) and 63 (PO_2^-). These phosphate specific fragment ions serve as characteristic "reporter ions" for phosphorylated peptides in several instrumental methods. These diagnostic ions can be selectively detected in "precursor-ion scanning" experiments by MS-MS. Triple quadrupole mass spectrometers are generally used for this purpose. In this scan type Q1 is scanned across mass range and ions are fragmented in collision cell (q2). Q3 is set to transmit only the mass of the diagnostic fragment . Therefore only ions that are passed through Q1 and produce the diagnostic charged fragment will be detected (fig.1).

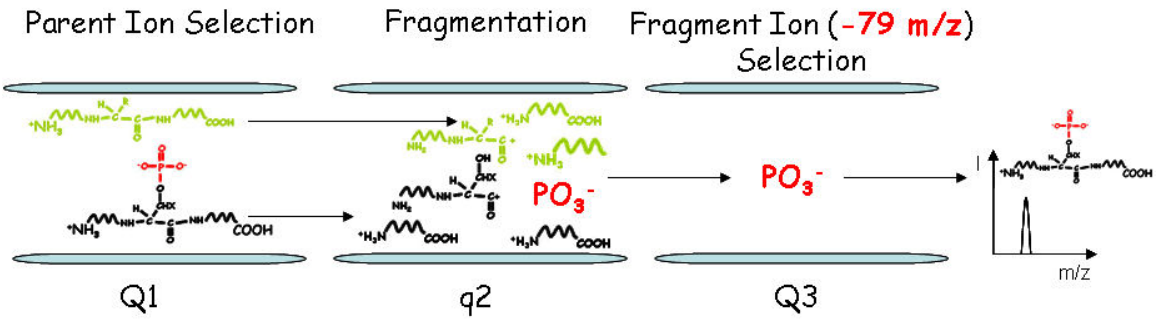


Figure 1. Precursor ion scan experiment.

The phosphopeptide analysis by precursor ion scans present the peculiarity that only those peptides that fragment producing the chosen marker ion, m/z 79, give rise to peaks in the spectrum, screening out all other species.

The following sequencing of the corresponding phosphopeptide precursors requires a change in polarity and in buffer of the sample in order to obtain good MS-MS spectra. This interferes with automation of the system. Despite this, the method is a powerful tool due to its high sensitivity and selectivity and its applicability for analysis of serine, threonine and tyrosine phosphorylated residues.

A precursor ion scanning method that can be performed in positive mode has recently been developed for specific detection of phosphotyrosine-containing peptides (5-6). This method is based on the ability to selectively detect the immonium ions of phosphotyrosine residues that have an m/z value of 216.04. Immonium ions are generated by double fragmentation of the peptide backbone (7).

CID of phosphoserine and phosphothreonine containing peptides in the positive ion mode often yields a neutral loss of H_3PO_4 via gas phase β-elimination reaction (8). Peaks corresponding to this loss (98 Da from singly charged precursor,

49 Da from doubly charged precursor, 32.66 from triply charged precursor, etc) are often the most abundant ions in the fragmentation spectrum.
In neutral loss experiment, Q1 is scanned across a specific mass range. Ions pass in collision cell where are fragmented. Q3 is scanned over a similar mass range, offset by the neutral mass of the diagnostic fragment. Therefore, any molecule passing through Q1 and losing a neutral molecule of the defined mass, will then be transmitted through Q3 and detected (fig.2).

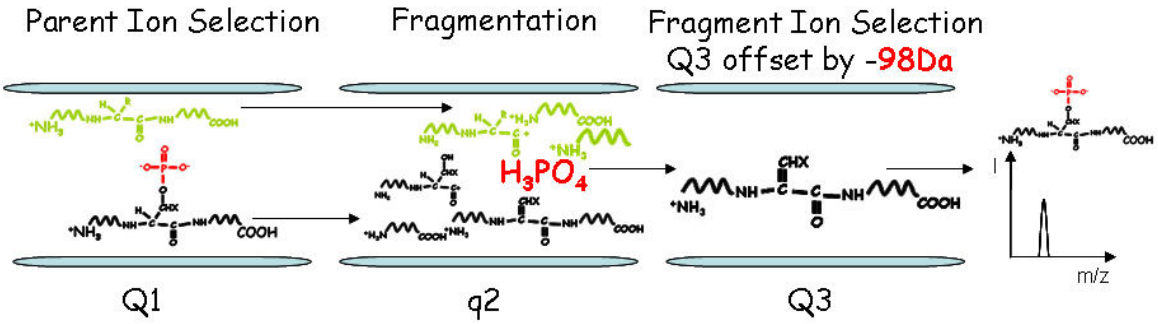


Figure 2. Neutral loss experiment.

Although mass spectrometry-based methods, such as precursor ion and neutral loss scans, are extremely useful for identifying unknown phosphopeptides from a complex mixture, they do not take advantage of any prior knowledge about the protein under investigation.

Quite often a significant amount of information is available. This may include the primary sequence, the type of phosphorylation (serine/threonine vs. tyrosine), or predicted phosphorylation sites.

This information can be used to predict precursor and fragment ion m/z values for a multiple reaction monitoring (MRM) experiment. MRM experiments, using a triple quadrupole instrument, are designed for obtaining the maximum sensitivity for detection of target compounds (8).
MRM is a highly sensitive function of triple quadrupole mass spectrometers where the first quadrupole transmits only an ion of specific m/z (Q1 mass). This ion is then fragmented in the second quadrupole. The third quadrupole is set to transmit a specific product ion (Q3 mass). Detection of this product is therefore diagnostic with multiple “Q1 to Q3 transitions” assayed in a single experiment. Only ions with this exact transition will be detected (fig.3).

Many MRM scans can be looped together in a single experiment to detect the presence of many specific ions in a complex mixture.

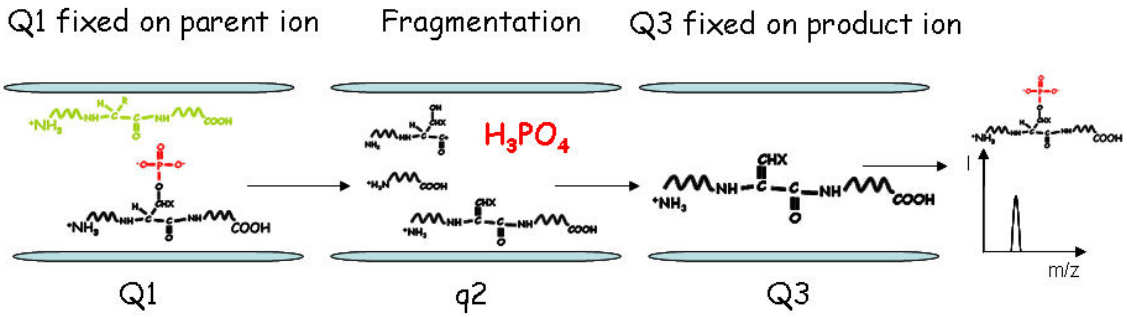


Figure 3. Multiple Reaction Monitoring experiment.

III.3 Peptidomics of human body fluids as sources for clinical marker

Searching for biomarker (specific proteins or peptides indicating the presents of abnormal states) in clinical diagnosis and therapy is one of the efforts in proteomics.

Human body fluids have been rediscovered in the post-genomic era as great source of biological markers and perhaps particularly as source of the potential biomarker of disease.

In particular, analysis of peptides in biological fluids holds promise of providing diagnostic and prognostic information for cancer and other diseases. “Peptidomics” is the field that deals with the comprehensive qualitative and quantitative analysis of peptides in biological samples (9). These peptides are either intact small molecules, such as hormomes, cytokines, growth factors or peptides that are released from larger protein precursors during protein processing. They may represent degradation products of proteolytic activity (fig 4).

Thus, in biological fluids, peptides represent protein synthesis, processing, and degradation. Since the amount and repertoire of peptides in the circulation change dynamically according to the physiological or pathological state of an individual, it is possible that comprehensive peptide analysis (i.e., exploitation of the “peptidome”) may lead to the discovery of novel biomarkers or to new diagnostic approaches (10).

With the development of new proteomic based tools to analyze tissue and blood samples from cancer patients, recently, many groups have reported the identification of biomarkers or biomarker patterns which allow the identification and detection of cancer patients (12) Among others, the surface-enhanced laser desorption ionization/time-of-flight mass spectrometry (SELDI/TOF-MS) has been used to identify differential peptide and protein expression in various biological fluids, such as serum, plasma, urine and pancreatic juice (13).

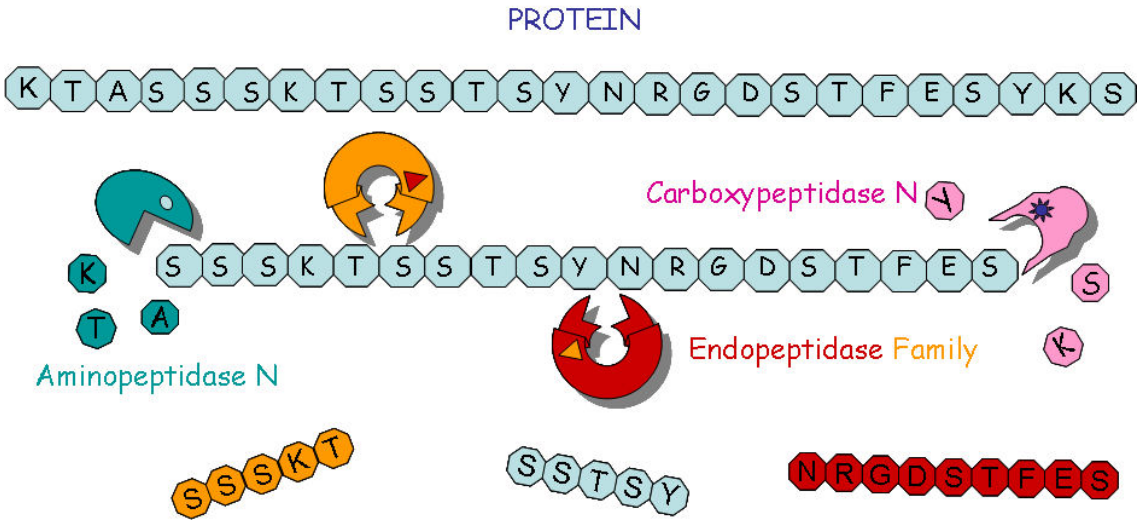


Figure 4. Mechanisms of generation of multiple peptides from a protein, used as an example. Initially, this fragment is proteolyzed by endoproteases. Subsequently, the original fragment and other fragments are further processed by aminopeptidases and carboxypeptidases, as shown. At the end, a large family of peptides is generated from the original fragment.

III.4 Aim of the project

Here, IMAC strategy has been optimized and coupled to iterative mass spectrometry based scanning techniques (neutral loss, precursor ion, multiple reactions monitoring) to improve detection of phosphorylated peptides.

This strategy was implemented and validated by using phosphorylated bovine α-casein and then applied to the study of the phosphopeptidomics in biologic fluids (serum, saliva, urine etc). In fact, biologic fluid phosphopeptides have recently gained increasing interest as potential source of the diagnostic marker for diseases. Furthermore, enriched tryptic α-casein phosphopeptides were used to set up the best conditions for the MRM scan mode and then used for the selective identification of phosphorylation sites of nm23H1, a protein involved in mammalian cancer.

III.5 Materials and Methods

Proteomic grade trypsin, α-casein, dithiothreitol, $FeCl_3$ and alfa-cyano-4-hydroxycinnamic acid were purchased from Sigma (St. Louis, MO, USA). Chelating Sepharose (IDA) (Amersham bioscience). Acetonitrile HPLC ultra gradient was from Romil. All used solvents were of the highest purity available from Baker (Phillipsburg, NJ). All other reagents and proteins were of the highest purity available from Sigma.

Reduction and alkylation of cysteine residues

Protein was reduced in 300mM Tris-HCl, 1.25 mM EDTA, pH 8.5 containing 6 M guanidinium chloride by incubation with a 10:1 molar excess of dithiothreitol over the total -SH groups at 37°C for two hours under nitrogen atmosphere. The free cysteine residues were alkylated by using a 5:1 molar excess over the total -SH groups of iodoacetamide, at room temperature for 30 min in the dark under nitrogen atmosphere. Protein samples were freed from the excess of reagents by gel filtration chromatography on PD10 Sephadex G25 column.

Enzymatic hydrolysis

Enzymatic digestion was carried out with trypsin in 50mM ammonium bicarbonate pH 8.5 at 37°C for 18 hours using an E/S ratio of 1/50 (w/w).

IMAC Cromatography

A ratio of 10:1 of peptide mixture to resin was used to ensure specific binding of phosphospecies. The resin IDA was charged with $FeCl_3$ and stored using the recommended procedure. The charged resin was equilibrated with binding solution (15 % acetic acid and 60 % Acetonitrile , pH 2.5). The dried peptide mixture was suspended with binding solution and incubated with charged resin for three hours at room temperature. The washes were performed using washing solution (60 acetonitrile, 5 % acetic acid) and a final wash by milly Q water. The eluition was performed using 2% NH_3 in water (pH 12).

Acetonitrile precipitation

Pathological serum samples were obtained from " Servizio Analisi del Secondo policlinico di Napoli". Salivary and urine samples were taken from 2 healthy voluntaries with informed consent.
2 volumes of HPLC ultra gradient Romil were added to 1 volume of fresh biologic fluids followed by gentle mixing for 30 min at room temperature. The sample were spun for 15 min at 12,000 rpm and the high abundant precipitated proteins were removed and the supernatant was concentrated in speed Vac system and lyophilized. Supernatat protein concentration was determined using a Bio-Rad protein assay according to manufacturer's instructions.

MALDI mass spectrometry

MALDI-TOF mass spectra were recorded using a Voyager DE-PRO mass spectrometer (Applied Biosystem, Framingham, USA) operating in reflector mode. A mixture of the peptide solution and α-cyano-hydroxycinnamic acid (10 mg/mL in 70% ACN and 10 mM citric acid in water) was applied to the metallic sample plate and dried at room temperature. The phosphopeptides mixture were mixed with 2,5-dihydroxybenzoic acid (25 mg/mL in 50% CAN and 2% phosphoric acid in water) and applied on the sample plate. Mass calibration was performed using a mixture of peptides from Applied Biosystem, containing des-Arg1-Bradykinin, Angiotensin I, Glu1-Fibrinopeptide B, ACTH (1-17), ACTH (18-39) and Insulin (bovine) as external standards. Raw data were analysed using Data Explorer software provided by Applied Biosystem and reported as monoisotopic masses.

nanoLC Mass Spectrometry

A mixture of peptide solution was analysed by LCMS analysis using a 4000Q-Trap (Applied Biosystems) coupled to an 1100 nano HPLC system (Agilent Technologies). The mixture was loaded on an Agilent reverse-phase pre-column cartridge (Zorbax 300 SB-C18, 5x0.3 mm, 5 μm) at 10 μl/min (A solvent 0.1% formic acid, loading time 5 min). Peptides were separated on a Agilent reverse-phase column (Zorbax 300 SB-C18, 150 mm X 75μm, 3.5 μm), at a flow rate of 0.3 μl/min with a 0% to 65% linear gradient in 60 min (A solvent 0.1% formic acid, 2% ACN in MQ water; B solvent 0.1% formic acid, 2% MQ water in ACN). Nanospray source was used at 2.5 kV with liquid coupling, with a declustering potential of 20 V, using an uncoated silica tip from NewObjectives (O.D. 150 µm, I.D. 20 µm, T.D. 10 µm). Data were acquired in information-dependent acquisition (IDA) mode, in which a full scan mass spectrum was followed by MS/MS of the 5 most abundant ions (2 s each). In particular, spectra acquisition of MS-MS analysis was based on a survey Enhanced MS Scan.(EMS) from 400 m/z to 1400 m/z at 4000 amu/sec. This scan mode was followed by an Enhanced Resolution experiment (ER) for the five most intense ions and then MS^2 spectra (EPI) were acquired using the best collision energy calculated on the bases of m/z values and charge state (rolling collision energy) from 100 m/z to 1400 m/z at 4000 amu/sec. Data were acquired and processed using Analyst software (Applied Biosystems).

Precursor ion scan

Spectra acquisition was based on a survey Precursor Ion Scanning. It was performed over a mass range of m/z 400-1400 at 333 amu/s (with Q1 set to low resolution and Q3 set to unit resolution) with a nanopray voltage of -2350 V applied to a Picotip New Objective (O.D. 150 µm, I.D. 50 µm, T.D. 15 µm) (Woburn,MA). Precursors were collided in Q2 with a collision energy ramp of -25 to -65 V across the mass range. If a precursor of -79 was detected above a pre-set threshold value (2500 cps), the polarity was automatically switched to +2500V (after a 700 ms dwell at 0V) with the same gas settings and a positive ion Enhanced Resolution Scan was performed at 250amu/s to determine the charge state of the ion. The Information dependent acquisition (IDA) software automatically adjusts the mass for the polarity switch from negative ion to positive ion, and assigns a more accurate monoisotopic mass.

Enhanced Product Ion (EPI) scans (MS/MS) were performed at 4000 amu/s and collision voltages were calculated automatically by rolling collision energy and it performed a maximum of one repeat before adding ion to the exclusion list for 60 s. Once this duty cycle was completed, the polarity was switched back to negative (after a 700 ms dwell at 0V) and the cycle repeated. The entire cycle duration, including fill times and processing times was less than 5.3 s.

Neutral loss

Spectra acquisition was based on a survey Neutral loss scanning. It was performed over a mass range of m/z 400-1400 at 333 amu/s (with Q1 set to low resolution and Q3 set to unit resolution) with a nanopray voltage of 2500 V applied to a Picotip New Objective (O.D. 150 µm, I.D. 50 µm, T.D. 15 µm) (Woburn,MA). Ions are passed in Q2 where are fragmented with a collision energy ramp of 20 to 50 V across the mass range. Q3 is scanned over a similar mass range, off set by the 49 Da. Therefore, the two most intense peaks that loss 49 Da are automatically selected for an ER and EPI scans with the above conditions.

Multiple reaction monitoring

4000 QTRAP in MRM mode for selective detection of phosphopeptides of α-casein and nm23H1 was operated (Multiple Reaction Monitoring).
The MRM transitions of potential α-casein and nm23H1 phosphopeptides were calculated by a software script developed by Applied Biosystems.

The time taken to analyze each MRM transition is termed the “dwell time.” These MRM transitions (50 msec dwell time) were used to trigger dependent linear ion trap scans: enhanced resolution and enhanced product ion (EPI) scans. The total cycle time for this method was 3–5 sec. This data-dependent method is referred to as targeted MRM-IDA, for multiple reaction-monitoring information-dependent acquisition.

In general, transitions were included for all tryptic peptides (maximum one missed cleavage) containing Ser, Thr, or Tyr residues with either one or two modifications and for doubly and triply charged species for the Q1 mass range 400–1600 m/z. The number of MRM transition is dependent on various factors, including protein size, number of peptides following digestion, and the number of potential phosphorylation sites. It is, however, important to optimize the cycle time to around or below 5 s. Such a cycle time ensures that if the peak width is around 0.5 min it is

highly probable that a peptide is scanned for and analyzed at least twice as it is eluted and that one of these analyses will occur at, or close to, the apex of its elution profile. The software requires an amino acid sequence of the protein of interest, a starter method containing the LC conditions and an empty MRM-IDA experiment. The software will perform an *in-silico* digest of the protein and create a set of peptides each containing at least one possible site of modification. For each peptide, it will generate an MRM transition for the calculated m/z of the precursor ion and an appropriate fragment ion. The new method, specific for the protein of interest, is saved and submitted as a batch for data acquisition. The Agilent Nanoflow LC system and 4000 QTRAP were both controlled using Analyst 1.4.1.

MASCOT analysis

Spectral data were analyzed using Analyst software (version 1.4.1) and MS-MS centroid peak lists were generated using the MASCOT.dll script (version 1.6b9). MS-MS centroid peaks were threshold at 0.1% of the base peak. MS/MS spectra were searched against Swiss Prot database using the licensed version of Mascot (Matrix Science), after converting the acquired MS-MS spectra in mascot generic file format. The Mascot search parameters were: taxonomy human for biological fluids and mammalian for α-casein; trypsin as enzyme for α-casein and nm23H1 and none for others cases allowing up to 3 missed cleavages, none fixed modifications, oxidation of M, pyroGlu N-term Q and phosphorylation of Y, S, T, as variable modifications, 0.4 Da MS/MS tolerance and 0.6 Da peptide tolerance. Spectral data were manually validated and contained sufficient information to assign not only the sequence, but also the site of phosphorylation. This conservative criterion resulted in omission of many peptides identified as being phosphorylated but for which spectral quality was insufficient to allow further characterization.

III.6 Results and discussion

This work described both the improvements to traditional IMAC and the use of a combination of fractionation procedures and complementary mass spectrometry based scanning approach to investigate the free phosphopeptides in biological fluids.

The validation of peptide IMAC enrichment of phosphorylated bovine α-casein tryptic peptides was achieved by MALDI-MS. Thus, an aliquot of α-casein was used to spike human serum prior to start the enrichment/analysis procedure.

The enriched phosphopeptide mixture was used to find the best conditions for the mass spectrometry based scanning techniques (neutral loss, precursor ion, multiple reactions monitoring).

III.6.1 Peptide IMAC sample preparation

Peptide IMAC was carried out with some modifications. Initial optimization of procedure was performed using tryptic peptides originating from α-casein. Commercial α-casein consists of α-casein S1 and α-casein S2 in traces and the preparations are usually contaminated with traces of beta-casein. Figure 5 shows

the sequence of α-casein. A list of the theoretical tryptic phosphorylated peptides derived from α-caseins S1-S2 and their molecular masses, is shown in table1

RPKHPIKHQGLPQEVLNENLLRFFVAPFPEVFGKEKVNELSKDIGSESTEDQAMEDIKQ
MEAESISSSEEIVPNSVEQKHIQKEDVPSERYLGYLEQLLRLKKYKVPQLEIVPNSAEER
LHSMKEGIHAQQKEPMIGVNQELAYFYPELFRQFYQLDAYPSGAWYYVPLGTQYTDAP
SFSDIPNPIGSENSEKTTMPLW

KNTMEHVSSSEESIISQETYKQEKNMAINPSKENLCSTFCKEVVRNANEEEYSIGSSSE
ESAEVATEEVKITVDDKHYQKALNEINQFYQKFPQYLQYLYQGPIVLNPWDQVKRNAV
PITPTLNREQLSTSEENSKKTVDMESTEVFTKKTKLTEEEKNRLNFLKKISQRYQKFALP
QYLKTVYQHQKAMKPWIQPKTKVIPYVRYL

Figure 5. α**-casein S1 and S2 sequence**.

Table 1. Phosphopeptides of bovine α-casein digested with trypsin.

Protein	Sequence	Monoisotopic mass	Modifications	Peptide
α-casein s1	DIGSESTEDQAMEDIK	1926.68	2P	43-58
α-casein s1	QMEAESISSSEEIVPNSVEQK	2719.90	5P	59-79
α-casein s1	KYKVVPQLEIVPNSAEER	2079.90	1P	103-119
α-casein s1	YKVPQLEIVPNSAEER	1950.94	1P	104-119
α-casein s1	VPQLEIVPNSAEER	1659.78	1P	106-119
α-casein s2	KNTMEHVSSSEESIISQETYK	2617.89	4P	1-21
α-casein s2	NANEEEYSIGSSSEESAEVATEEVK	3007.72	4P	46-70
α-casein s2	EQLSTSEENSK	1410.49	2P	126-136
α-casein s2	TVDMESTEVFTK	1465.60	1P	138-149

Evaluation of the phosphorylated peptide binding selectivity of the IMAC resin and the optimization of the binding/washing and eluting conditions was performed by comparing the relative intensities of the phosphorylated/non-phosphorylated tryptic peptides.

A number of conditions in this protocol can affect the yield and population of the obtained phosphopeptides. The nature of the chelator, the stationary phase and the buffer systems are factors potentially important in improving selectivity. Two polysaccharide-based matrix were tested: sepharose with imminodiacetic acid (IDA) (Amersham bioscience) a tridentate ligand and agarose with nitrilotriacetic acid (NTA) (Qiagen) a quadradentate ligand (fig. 6).

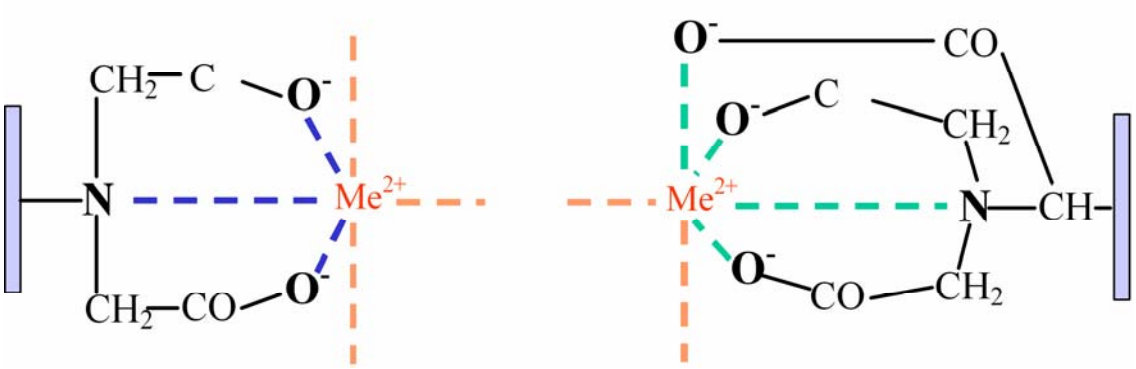


Figure 6. Structure of IDA and NTA ligands.

These chelating groups were charged with $FeCl_3$ using the recommended procedure.

The effectiveness of the isolation procedure was tested by using an equal amount of α-casein peptide mixture under standard conditions. The eluted phosphopeptides were analyzed by MALDI-MS to verify the enrichment protocol. Then, enriched peptide mixture was mixed in a ratio 1:1 (v/v) with a matrix solution of 2,5-DHB with phosphoric acid (see materials and methods section) in order, to improve the phosphopeptides ion response in MALDI-MS (17).

The phosphopeptides were identified by their 80 Da mass difference compared to the peptides expected from the sequence for the presence of the phosphate moiety. As an example, the signal at m/z 1660.6 was assigned to the peptide 104-119 within α-casein sequence carrying a phosphate group as reported in figure 7.

As clearly indicated in figure 7, enrichment procedure performed by using Fe^{3+}-IDA-agarose (panel A) showed greater selectivity and recovery in binding phosphopeptides versus non phosphorylated peptides, compared to Fe^{3+}-NTA-agarose (panel B) and therefore it was used in all subsequent IMAC experiments.

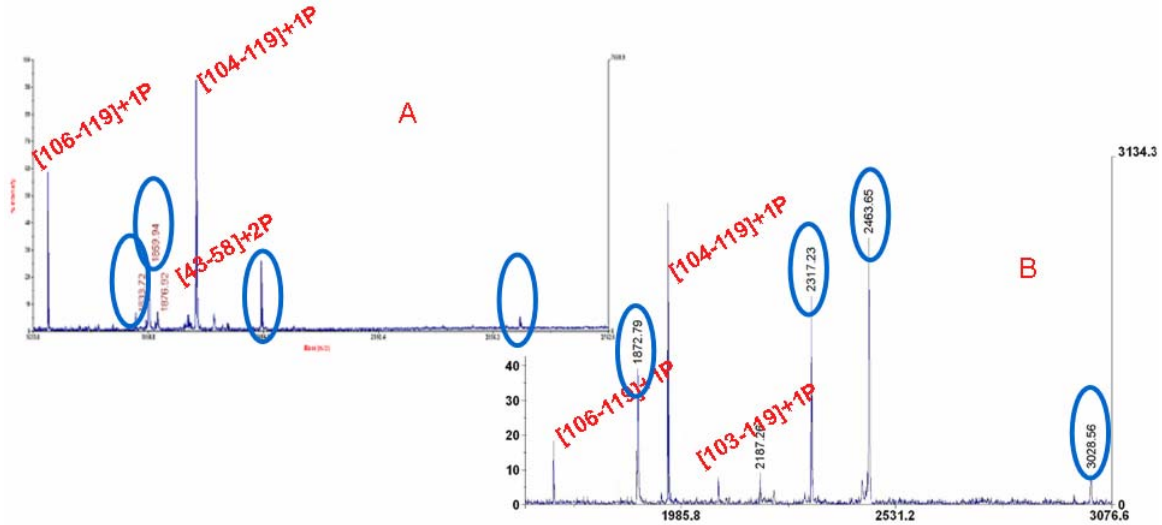


Figure 7 A MALDI-MS analysis of enriched α-casein by Fe^{3+}-IDA-agarose; B MALDI-MS of enriched α-casein by Fe^{3+}-NTA-agarose. In blue are indicated unspecific no phosphorylated peptides.

It should be considered that in these conditions the pentaphosphorylated peptide occurring within α-casein sequence was not detected.

Acidic peptides affect the specificity of peptide IMAC purification. To minimize the effect of non specific binding, methyl-esterification of peptide side chains was performed on the peptide mixture as suggested by Ficarro *et al* (37).

In this method, the C-termini and aspartic and glutamic acid residues of the tryptic digest peptides were methylated prior to IMAC in order to preclude the previously observed non-specific binding. As shown in figure 8, MALDI-MS analysis revealed the presence of S1 α-casein phosphopeptides. It is worth considering that in the high mass region of spectrum a signal at m/z 2749.12 was detected. This signal was attributed to the peptide 59-79 carrying up to five phosphate moieties. Moreover unphosphorylated peptides persisted, possibly due to unspecific binding to the resin.

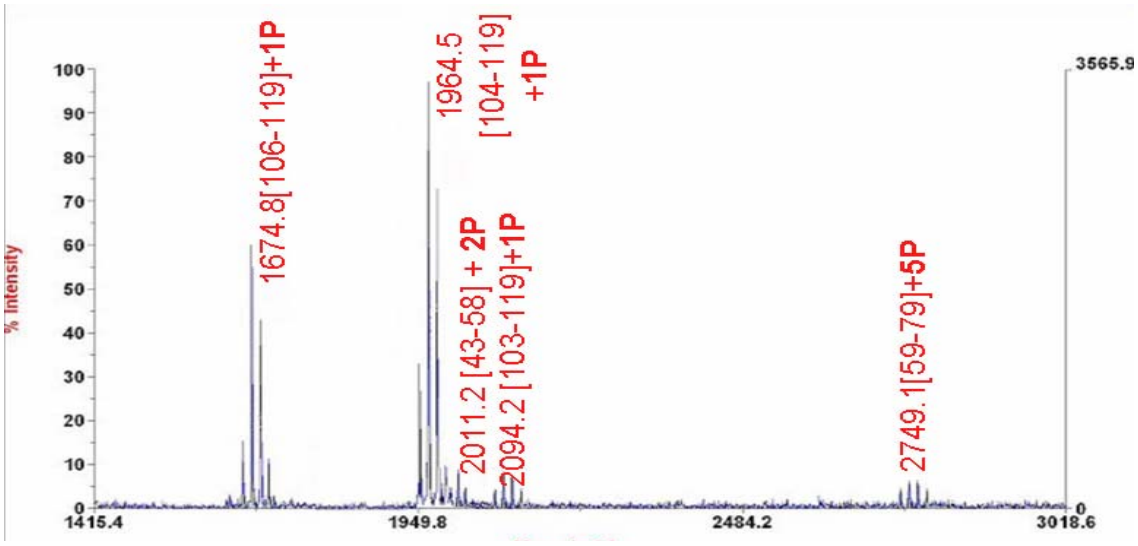


Figure 8. MALDI-MS analysis of enriched methylated α-casein.

However, peptide chemical modification reaction was not quantitative, as indicated by the occurrence of the satellite peaks at 14 Da higher than the theoretical ones. This effect diluted the sample into populations of modified and unmodified peptides. This effectively reduced the amount of any given peptide and increased the overall sample complexity. For this reason, methylation was omitted in the present study.

To minimize non specific binding in IMAC step, the pH of sample was carefully adjusted to 2.5 value, which resulted in the protonation of the majority of glutamic and aspartic acid residues. The pKa_1 value of the phosphoric acid is 1.8 and therefore the phosphate group will still have had a negative charge at pH 2.5. In addition, a high-organic (60 % acetonitrile, 15% acetic acid pH 2.5) buffer was used to minimize the hydrophobic unspecific interactions with matrix and solubilize peptides before loading onto the IMAC resin. The elution was performed by pH changing. As eluting buffer a solution of volatile ammonia (pH 12) was chosen thus avoiding the presence of salts and other contaminants.

The feasibility of the developed strategy to detect phosphopeptides in a highly complex mixture was probed by adding an aliquot of α-casein to human serum.

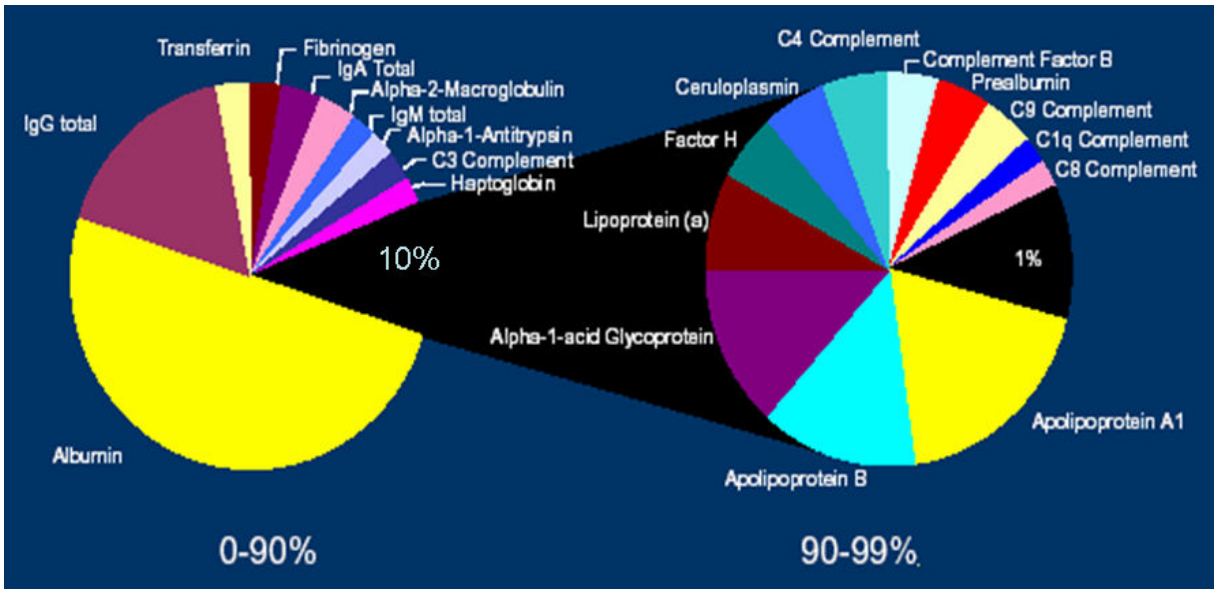


Figure 9. Protein composition of human serum.

Human serum is a complex body fluid that likely contains tens of thousands of proteins and peptides as shown in figure 9. Serum contains 60-80 mg/ml of proteins in addition to various small molecules including salts, amino acid and sugars. The major protein constituents of serum include albumin, immunoglobulins, trasferrin, and others.

After enzymatic digestion, the complexity of mixture was increased of two orders magnitude (about 10^8 peptide).

As shown in figure 10 and table 2, mass spectral analysis led to a complete coverage of S1 α-casein phosphopeptides and partial coverage of S2 phosphopeptides. Only few signals attributed to non phosphorylated peptides were detected thus demonstrating the success of the enrichment strategy.

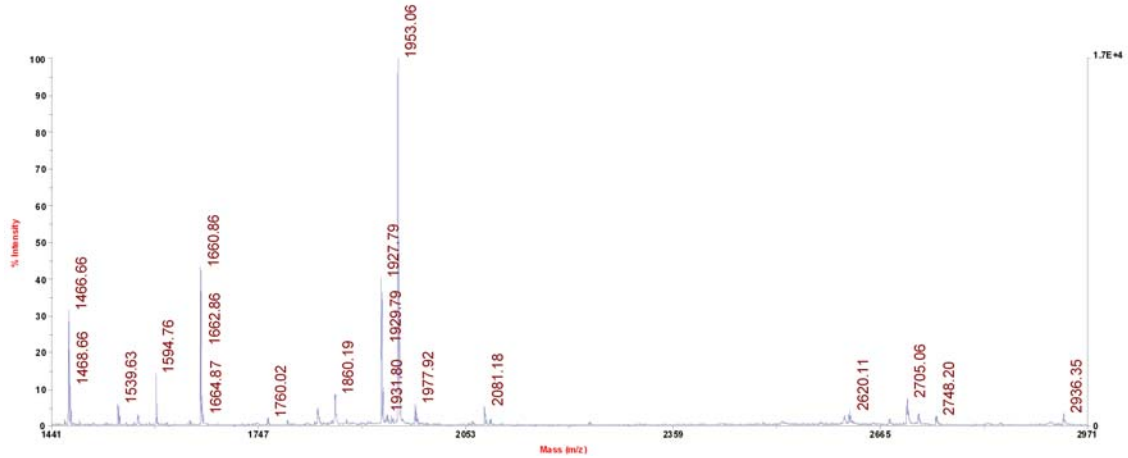


Figure10. MALDI-MS analysis of enriched α-casein mixture in serum sample.

Table 2. List of identified α-casein phosphopeptides by MALDI-MS analysis.

Protein	Sequence	Observed m/z	Modifications	Peptide
α-casein s1	DIGSESTEDQAMEDIK	1927.79	2P	43-58
α-casein s1	QMEAESISSSEEIVPNSVEQK	2721.14	5P	59-79
α-casein s1	KYKVVPQLEIVPNSAEER	2080.18	1P	103-119
α-casein s1	YKVPQLEIVPNSAEER	1952.06	1P	104-119
α-casein s1	VPQLEIVPNSAEER	1660.86	1P	106-119
α-casein s2	NTMEHVSSSEESIISQETYK	2619.09	4P	2-21
α-casein s2	KNTMEHVSSSEESIISQETYK	2747.20	4P	1-21
α-casein s2	TVDMESTEVFTKk	1594.76	1P	138-150
α-casein s2	TVDMESTEVFTK	1466.66	1P	138-149

III.6.2 Unique scanning capabilities using a Q-q-Q linear ion trap (Q TRAP) mass spectrometer

Mapping phosphorylation sites within phosphoproteins by MALDI-MS is currently the method of choice for identifying phosphorylation sites because of its high sensitivity and accuracy. However, MALDI-MS analysis of phosphopeptides is challenging due to low ionization efficiency and suppression effects in the MS positive mode. Furthermore, the MALDI mapping analysis of the phosphopeptides and the identification of the phosphorylation sites is hard task without an *a priori* knowledge of protein sequence thus avoiding an application of phosphoproteomic analysis on a large scale. In addition, the presence of several possible phosphorylation sites complicates the correct assignment of phosphorylated residues.

Reducing the complexity of peptide mixtures by enriching and separating phosphopeptides by liquid chromatography is the common method used in phosphoproteomics. Coupling liquid chromatography to hybrid mass spectrometers (LC-MS-MS) can overcome above problems leading to both the identification of phosphoproteins and the exact location of phosphorylation sites.

In this work the detection of phosphorylated peptides was improved by selectively monitoring phosphopeptides in the mass spectrometer taking advantage of scanning capabilities of a new generation of instruments based on linear ion traps.

The recently developed “linear ion traps” (LIT) offer faster scanning procedures and higher ion capacities due to larger trapping volumes together with the possibility of being used as third quadrupole in a triple quadrupole instrument. Thereby, the benefits of triple quadrupole instruments and ion traps are combined.

The work described here was carried out on a 4000 Q-Trap (Applied Biosystems). Specific mass spectrometry scanning modes, such as neutral loss, precursor ion scanning and multiple reaction monitoring (MRM), have been implemented to improve the detection of the phosphorylation sites in highly complex peptide mixture. Furthermore, the nanospray ESI source combined with capillary chromatography system showed performances characterized by higher sensitivity and resolution, leading to the detection of samples even present at substoichiometric amount as phosphopeptides.

In order to evaluate the performances of the LIT system and to determine the best acquisition conditions for neutral loss, precursor ion scanning and MRM,

standard samples of enriched α-casein phosphopeptides were used in preliminary experiments.

Precursor ion scan of α-casein

Precursor ion scan can be used to selectively identify phosphorylated peptides by detecting the m/z -79 (PO_3^-) product ions released during MS-MS.

Briefly, in the precursor survey scan Q1 was set to low resolution and Q3 was operated at unit resolution to increase selectivity while keeping sensitivity. This step was performed in the negative ion mode. Once the ion of interest was detected in the survey scan, the instrument was toggled to positive mode and an Enhanced Resolution (ER) scan is performed. This scan was automatically performed at the highest resolution available, in order to improve mass accuracy on the ^{12}C isotope and consequentially confirm the charge state. The cycle was completed with an EPI scan in the positive ion mode to provide fragmentation information for sequence tagging and database searching.

The precursor ion analysis of tryptic α-casein peptides before the affinity chromatography identified only monophosphorylated phosphopeptides within S1 α-casein sequence.

The enriched α-casein mixture was submitted to the same selective mass spectrometry analysis.

In this case, a large number of phosphorylated peptide ions were detected within S1 α-casein and S2 α-casein sequence. Table 3 summarized mass spectra results. The known phosphorylation sites confirmed by mass spectral data were indicated in yellow; a new phosphorylation site was indicated in red. These data emphasised the importance of the enrichment step.

Table 3. Identified α-casein phosphopeptides by precursor ion scan analysis.

Protein	Observed m/z	Monoisotopic mass	Sequence	Phosphorylation site
α-casein s1	554.30	1659.79	VPQLEIVPNSAEER	S115
α-casein s1	651.27-976.61	1950.95	YKVPQLEIVPNSAEER	S115
α-casein s1	643.2-963.36	1926.68	DIGSESTEDQAMEDIK	S46-S48
α-casein s2	670.21	2676.83	VNELSKDIGSESTEDQAMEDIK	S41-S46-S48
α-casein s2	733.81	1465.6	TVDMESTEVFTK	S143

As an example**,** figure 11 reports the identification steps of phosphopeptides performed on enriched α-casein trypsin digest.

Panel A shows the survey scan of a signal at 641.2 m/z and at the retention time of 30.7 min. The analysis of phosphopeptides was restricted only to those species whose fragmentation produces the marker ion at m/z -79 due to the presence to the phosphate moiety. The specificity of the precursor scan was clearly indicated by the simplicity of the survey MS spectrum.

Having proved by the precursor ion scan that the peptide at m/z 641.2 was phosphorylated, in order to fragment it the following step was to determine the exact charge state. The resulting ER scan performed in the positive mode for the same peptide was reported in panel B, thus indicating that the corresponding precursor ion at m/z 641.2 was a triply charged ion.

The fine localization of the phosphorylation sites in the peptide was achieved by MS-MS sequencing. The corresponding MS^2 scan was shown in Figure 11 C.

The MS-MS spectrum revealed the occurrence of the complete series of C-terminal fragment ions y1-y14 for the modified phosphopeptide carrying out two phosphate groups on serine residues.

No pentaphosphorylated peptide was detected. This was due to the worst ionization of this peptide in *electrospray* positive mode.

The selectivity of precursor ion scanning for phosphopeptides has been reported previously (18-19-20).

However, there are a few cases where a precursor of 79 signal will be generated from non-phosphorylated peptides, most commonly from peptides containing cysteic acid and methionine sulphone. Therefore oxidation of both cysteine and methionine by performic acid should be avoided, if this PTM discovery workflow is to be undertaken.

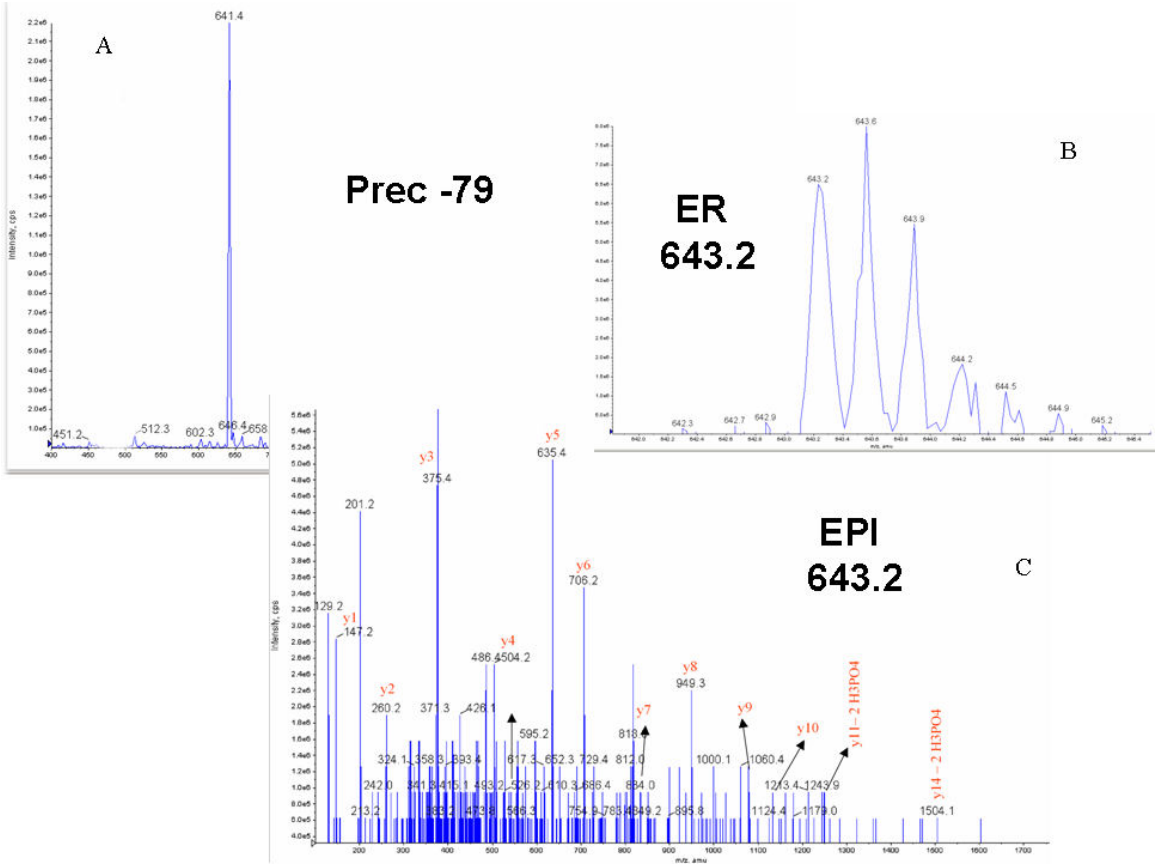


Figure 11. Precursor ion scan analysis of the diphosphorylated triply charged peptide at m/z 641.2.

Neutral loss scan of α-casein

The potential of a neutral loss scan was investigated. Though a high degree of specificity is achieved by precursor ion scan, the duty cycle of the experiment is reduced, mainly due to the switching of polarity. In order to reduce the cycle time, a neutral loss in the positive mode of 98, 49 or 32.7 (corresponding to the neutral loss of phosphoric acid from 1+, 2+, and 3+ charge states, respectively) was used as an alternative for the survey scan. This approach eliminates the overheads associated with the switching of instrument polarity and provides a gain of 1.4 seconds in the cycle time.

A mixture of enriched IMAC α-casein peptides, was submitted to neutral loss analysis. The analysis was performed using the neutral loss of 49 as survey scan and the enhanced MS (EMS) as comparison.

The spectra of phosphopeptide 104-119 within α-casein sequence related to the survey experiment steps are shown in figure 12. The selectivity of the neutral loss experiment is clearly evident. In fact, the corresponding trace shows the detection of only one relevant peptide at m/z 976.5, whereas the EMS trace, though more sensitive by 1–2 orders of magnitude, is more complex as illustrated in figure 12A and 12B respectively.

Traditional ion trap single EMS "survey scans" can be very sensitive, but they lack the selectivity required to identify peculiar peptide species in complex matrix. Furthermore, the complexity of the EMS spectrum and the relative intensity between phosphorylated/non phosphorylated species illustrated the challenge of the phosphoproteome analysis by mass spectrometry. However, at the same retention time, the neutral loss scan showed only one peak of interest, thus greatly simplifying the analysis.

The loss of isotope information, and therefore charge state assignment, from the neutral loss experiment was gained by performing an ER scan for that mass.

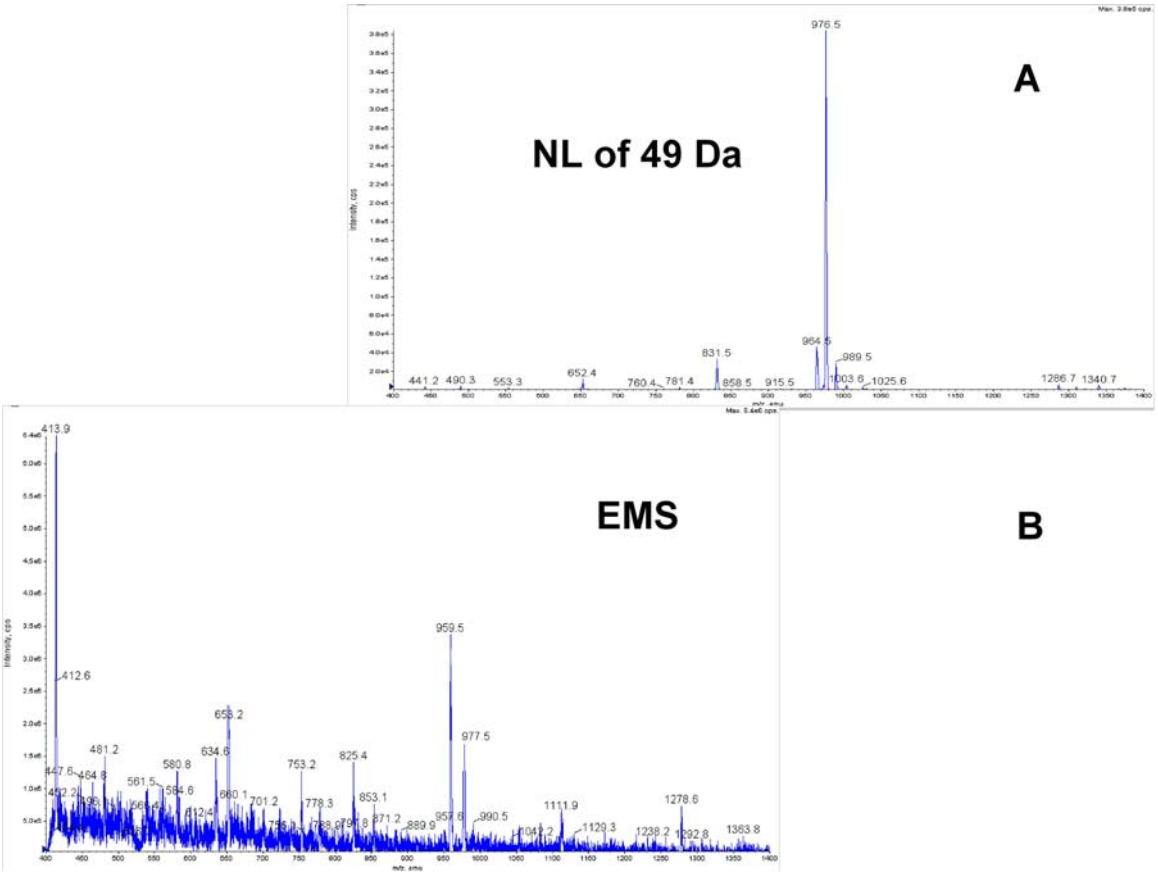


Figure 12. Comparison between the survey scan of Neutral Loss (A) and Enhanced MS EMS survey scan (B) at the retention time of 32.5. The EMS spectrum indicates that the ion of interest (m/z 976), though present, was suppressed by other species from the digest mixture that are co-eluting.

The specificity of neutral loss scan is due the collision energy to produce the phosphoric acid loss from peptide. In a spectrum generated by scanning for neutral loss of phosphoric acid the signal-to-noise ratio of phosphopeptides signals depends

on the abundance of the fragmentation reaction and on the abundance of interfering ion signals.

The amount of phosphoric acid loss was optimized by a standard phosphopeptide (monophosphorylated peptide of α-casein) and by selection of the optimal collision offset. With regard to interfering fragment signals, the ideal situation for application of neutral loss scanning would be the one in which the loss of phosphoric acid is the only or at least the most prominent fragmentation reaction. However, since peptides show an increasing tendency to backbone fragmentation with increasing charge state, there is a risk that multiply charged non phosphorylated peptides may show fragments located at -98/z Da relative to the molecular ion (z, number of charges) and thus giving rise to a signal in the neutral loss scan. The main reason for this effect is that the low collision offsets required for optimal detection of doubly and triply charged phosphopeptides are also optimal for peptide backbone fragmentation of multiply charged peptides.

In this protocol the best value of collision offset for the neutral loss scan was ramped from start to 20 Collision Energy Volt and end to 50 Collision Energy Volt to minimize side fragmentations.

This strategy was successful thanks to the detection of very few unphosphorylated peptides, it confirmed the identification performed by MALDI-MS and added new phosphopeptides of S1 α-casein phosphopeptides, S2 α-casein phosphopeptides and the beta casein one, as shown in table 4.

Table 4. Identified α-casein phosphopeptides by neutral loss analysis.

Protein	Observed m/z	Monoisotopic mass	Sequence	Phosphorylation site
α casein S1	830.97	1659.79	VPQLEIVPNSAEER	S 115
α casein S1	976.51	1950.95	YKVPQLEIVPNSAEER	S115
α casein S1	964.36	1926.68	DIGSESTEDQAMEDIK	S46-s48
α casein S1	972.36	1942.68	DIGSESTEDQAM(ox)EDIK	S46-s48
α casein S2	733.86	1465.71	TVDMESTEVFTK	S 143
α casein S2	741.82	1481.63	TVDM(ox)ESTEVFTK	S 143
α casein S2	797.84	1593.67	KTVDMESTEVFTK	S 143
α casein S2	805.87	1609.69	KTVDM(ox)ESTEVFTK	S 143
Beta casein	1031.44	2060.86	FQSEEQQQTEDELQDK	S35

MRM Analysis.

This method made use of MRM as a selective trigger for enhanced product ion (MS/MS) scans on phosphopeptides.
MRM requires two ions to generate a positive result, making it very specific with very low background, enhancing sensitivity of detection.

The Q1 mass is calculated for one peptide in doubly/triply charged form with up to two phosphate additions. The Q3 mass is calculated as either loss of 98 Da for each putative Ser(P) or Thr(P) or generation of y2 or b2 ions.

This differed significantly from the above mass spectrometric technique (neutral loss scan) for identifying phosphopeptides (21). In a neutral loss scan, Q1 scanned a specified mass range while Q3 scanned the same range, even if it lagged behind by a specified m/z (for example 49 m/z is useful for identifying doubly charged phosphopeptides). A neutral loss scan attempted to find any compound that

will lose a functional group during CID; while an MRM scan looked for a target peptide to generate a specific fragment. A common problem in using a neutral loss scan to discover phosphopeptides was that many phosphopeptides could generate a similar neutral loss, causing the mass spectrometer to waste time identifying a non-modified peptide. Or worse, a phosphopeptide will not generate the specified neutral loss, and therefore will not be detected. A neutral loss scan is limited to the investigation of only one predicted fragmentation pathway (e.g., loss of 49 m/z), while a targeted MRM approach can investigate several predicted fragmentation pathways for the same peptide (loss of 49 m/z from 2+, loss of 32.7 m/z from 3+, formation of a b2 ion, etc.) .

The sensitivity and selectivity of this method were assessed using the model phosphoprotein, α-casein. Two ions were monitored: loss of 98 Da and b2 ions. These analysis confirmed the previous attributions and also a new phosphorylation site within α-casein sequence. These results emphasized the high sensitivity of this experiment.
The results are summarised in tables 5 and 6.

Table. 5 MRM analysis of enriched α-casein peptides by monitoring b2 ions.

Observed Transition	Sequence	Phosphorylation site
Q1 830.8/Q3 197.1	VPQLEIVPNSAEER	S 115
Q1 554.3/Q3 197.1	VPQLEIVPNSAEER	S 115
Q1 976.4/Q3 292.1	YKVPQLEIVPNSAEER	S 115
Q1 651.2/Q3 292.1	YKVPQLEIVPNSAEER	S 115
Q1 964.8/Q3 229.1	DIGSESTEDQAMEDIK	S46-S48
Q1 643.8/Q3 229.1	DIGSESTEDQAMEDIK	S46-S48
Q1 648.5/Q3 229.1	DIGSESTEDQAM(ox)EDIK	S46-S48
Q1 972.3/Q3 229.1	DIGSESTEDQAM(ox)EDIK	S46-S48
Q1 866.6/Q3 214.1	VNELSKDIGSESTEDQAMEDIK	S46-S48
Q1 893.3/Q3 214.1	VNELSKDIGSESTEDQAMEDIK	S41-S46-S48

Table. 6 MRM analysis of enriched α-casein peptides by monitoring loss of 49 Da ions.

Observed Transition	Sequence	Phosphorylation site
Q1 830.8/Q3 781.8	VPQLEIVPNSAEER	S 115
Q1 554.3/Q3 521.5	VPQLEIVPNSAEER	S 115
Q1 976.4/Q3 927.4	YKVPQLEIVPNSAEER	S 115
Q1 651.3/Q3 618.6	YKVPQLEIVPNSAEER	S 115
Q1 964.8/Q3 915.3	DIGSESTEDQAMEDIK	S46-S48
Q1 923.3/Q3 875.3	DIGSESTEDQAMEDIK	S48
Q1 932.3/Q3 883.3	DIGSESTEDQAM(ox)EDIK	S48
Q1 972.3/Q3 923.3	DIGSESTEDQAM(ox)EDIK	S46-S48
Q1 866.6/Q3 834.0	VNELSKDIGSESTEDQAMEDIK	S46-S48
Q1 893.3/Q3 860.6	VNELSKDIGSESTEDQAMEDIK	S41-S46-S48

This strategy was applied to the characterization of the phosphorylation state of the nm23H1 (fig.13), a protein involved in tumor cell motility (38). The knowledge of the primary sequence, the type of phosphorylation on serine and the predicted phosphorilation sites allowed to predict precursor and fragment ion m/z values for a multiple reaction monitoring (MRM) experiment.

ANCERTFIAIKPDGVQRGLVGEIIKRFEQKGFRLVGLKFMQASEDLLKEHYVDLKDRPF
FAGLVKYMHSGPVVAMVWEGLNVVKTGRVMLGETNPADSKPGTIRGDFCIQVGRNIIH
GSDSVESAEKEIGLWFHPEELVDYTSCAQNWIYE

Figure 13. nm23-H1 sequence.

The nm23H1 protein samples were reduced, alkylated and digested with trypsin (see Material and Methods section). The peptide mixture was then directly analysed by MALDI-MS. No phosphopeptide was detected by this analysis. The phosphopeptides and the phosphorylation sites were determined by tandem mass spectrometry analysis and database searching. The enriched peptide mixture was then submitted to LC-MS analysis by using 4000 Q-Trap in MRM mode coupled to a nano-HPLC system.

A MRM experiment was performed. The Analyst software created a method with 45 transitions, corresponding to potential peptides phosphorylated on serine from nm23H1. Therefore, it became possible to probe for several predicted phosphopeptides from a known protein sequence.

For example, a method using a specific MRM transition was employed to detect a phosphopeptide within nm23H1 sequence. In fact, the theoretical phosphoserine-containing peptide 115-128 giving rise to a doubly charged ion at m/z 783.3, will lose a phosphate group (49 m/z) during CID to form a doubly charged fragment at m/z 734.3. These two values, Q1 = 783.3 and Q3 = 734.3, will be used to specifically detect this phosphopeptide.

The MS-MS spectrum performed on the precursor ion at m/z 783.3 assigned at peptide 115-128 NIIHGSDSVESAEK within nm23H1 sequence, indicated the phosphorylation site at Ser122. The same peptide 115-128 at m/z 823.27 doubly charged was identified carrying a phosphate group on Ser 122 and Ser 125. This information was important to understand the regulator mechanism of nm23H1 during tumour development.

Mascot analysis of the phosphopeptides

All fragmentation spectra were interpreted both using Mascot software and *de visu* inspection. The Mascot expectation values used to confirm individual phosphopeptides spectra did not directly address the probability that a given serine, threonine, or tyrosine residue within the peptide was the phosphorylation site. If a phosphopeptide has more than one potential modification site and its spectrum is analyzed using MASCOT, one can at best compare the MASCOT peptide scores associated with phosphorylation to each potential sites to discriminate the modified one. Furthermore, Mascot relies on neutral loss of phosphoric acid (which appears as a loss of 98 Da from the modified amino acid) in MS-MS spectra to designate the exact phosphorylation site. However, this strategy suffers of some limits. In fact, not phosphorylated peptide fragments can undergo loss of H_2O, yielding fragment ions with the same mass as a loss of H_3PO_4 from a peptide fragment ion containing a phosphorylated serine or threonine. Then, during fragmentation, the phosphate group on a modified serine or threonine residues does not always undergo complete neutral loss of H_3PO_4. Fragment ions in which the phosphate group is retained can provide valuable information regarding which amino acid is the modification site.

Finally, instead of the neutral loss of H_3PO_4, phosphorylated serine and threonine fragment ions can undergo the neutral loss of HPO_3. This can lead to the misidentification of the site of phosphorylation if these -80 Da fragments are misinterpreted as initially not phosphorylated sequence ions. All phosphopeptides were manually inspected to verify and complete the MASCOT assignments of y-H_20/H_3PO_4 or b-H_2O/H_3PO_4 ions. Furthermore many MS-MS spectra had not MASCOT interpretation and were assigned by *de novo* sequence analysis.

III.6.3 PhosphoPeptidomics of biological fluids

One goal of this study was to apply this strategy to the identification of free phosphopeptides in serum, saliva and urine.

These samples contain large amount of proteins (for example albumin, amylase and uromodulin respectively) that must be depleted before the analysis of less abundant proteins. Any complexity reduction strategy greatly increases the number of less abundant proteins that can be subsequently analyzed.

Biological fluids have been shown to contain peptides resulting from degradation of proteins and other tissues and the change in composition of these peptides and their PTMs can be indicative of changes in the physiology of organism. Because such peptides could be bound to highly abundant proteins, their depletion from biological fluids under native conditions may possibly result in the loss of the potential biomarkers.

A combination of a precipitation protocol of the higher-abundance proteins and IMAC strategy was chosen.

Acetonitrile Precipitation

Many proteins are known to act as carriers for smaller proteins or peptides. Techniques to remove these abundant molecules using affinity columns or filtration systems may result in the loss of these smaller proteins of interest. For this reason, acetonitrile, was chosen to precipitate the large proteins out of solution (see Material and Methods section). These conditions dissociated smaller proteins from their carrier molecules, allowing the detection of a larger number of peptides (22).

MS Analysis of phosphopeptide in biological fluids

The supernatants were dry and the free peptide mixture of biological fluids was submitted to MALDI-MS and LC-MS-MS analysis. MALDI-MS analysis shown in Figure 14 was carried out to know the complexity of the samples and the masses of their peptides. However, the MALDI mapping procedure was not sufficient to identify any peptide. Thus, the same mixture was submitted to LC-MS-MS analysis. Tandem mass spectrometry allowed to know the sequence of peptides and thus the identity of proteins in the mixture as reported in table 7. All fragmentation spectra were identified by MASCOT searching inserting also phosho-Ser and Thr among variable modifications. Either in the serum or saliva or urine no free phosphopeptides were detected.

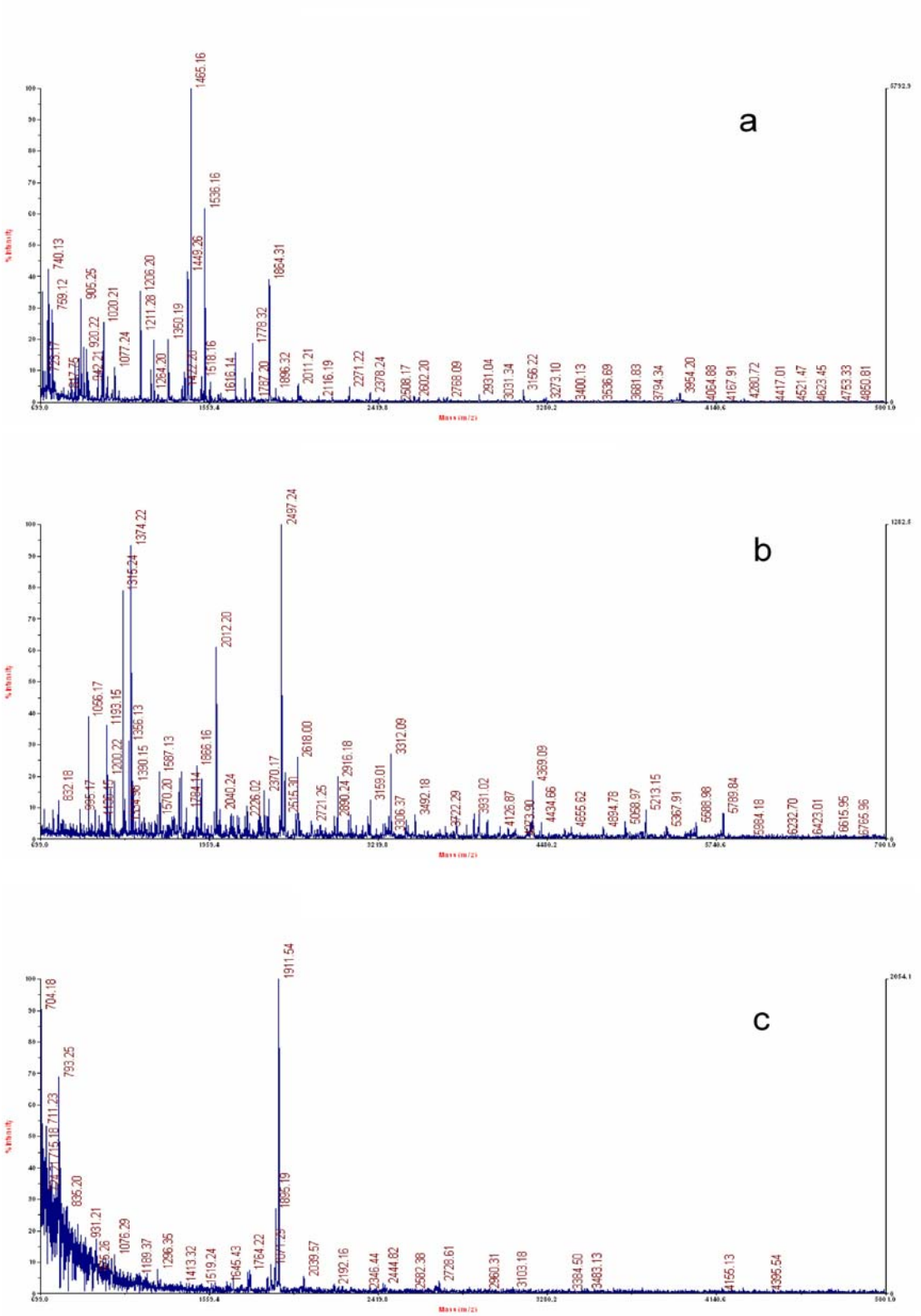


Figure. 14 a, b, c MALDI-MS analysis of precipitated peptides from serum, saliva and urine respectively.

Table 7. Identified proteins by MS-MS analysis of precipitated peptides from biological fluids.

Origin	Identified protein	Number of peptides
Saliva	(P02814) Submaxillary gland androgen-regulated protein 3 homolog B precursor (Proline-rich protein	19
Saliva	(P02812) Basic salivary proline-rich protein 2 (Salivary proline-rich protein) (Con1 glycoprotein)	30
Saliva	(P04281) Basic proline-rich peptide IB-1	4
Saliva	(P02811) Basic proline-rich peptide P-E (IB-9)	1
Saliva	(P02810) Salivary acidic proline-rich phosphoprotein 1/2 precursor (PRP-1/PRP-2) (Parotid proline-r	19
Saliva	(P10163) Basic salivary proline-rich protein 4 allele S precursor (Salivary proline-rich protein Po	16
Saliva	(P02808) Statherin precursor	2
Saliva	(Q14050) Collagen alpha-3(IX) chain precursor	3
Saliva	(Q9UMD9) Collagen alpha-1(XVII) chain (Bullous pemphigoid antigen 2) (180 kDa bullous pemphigoid an	2
Saliva	(Q04118) Basic salivary proline-rich protein 3 precursor (Parotid salivary glycoprotein G1) (Prolin	4
Urine	(P07911)Uromodulin precursor	7
Serum	P02671 Fibrinogen alpha chain precursor	15
Serum	P01024 Complement C3 precursor	3
Serum	P22670 MHC class II Enhancer factor C	1

IMAC protocol and selective monitoring of free phosphopeptides by mass spectrometry

Improved IMAC protocol was applied to enrich phosphopeptides mixture in order to identify the phosphorylated peptides inside biological fluids

The conditions set up for model phosphopeptides analysis were employed for the enriched free phosphopeptides in biologic fluids.

MALDI-MS analysis performed on serum and saliva after IMAC enrichment step was reported in figure 15 A and B respectively. MALDI-MS data, shown in panel A and B, clearly indicated the presence of a high number of signals with respect to the same analysis performed on the sample before IMAC enrichment (fig. 14 A-B). MALDI-MS analysis of enriched urine phosphopeptides did not give such good results.

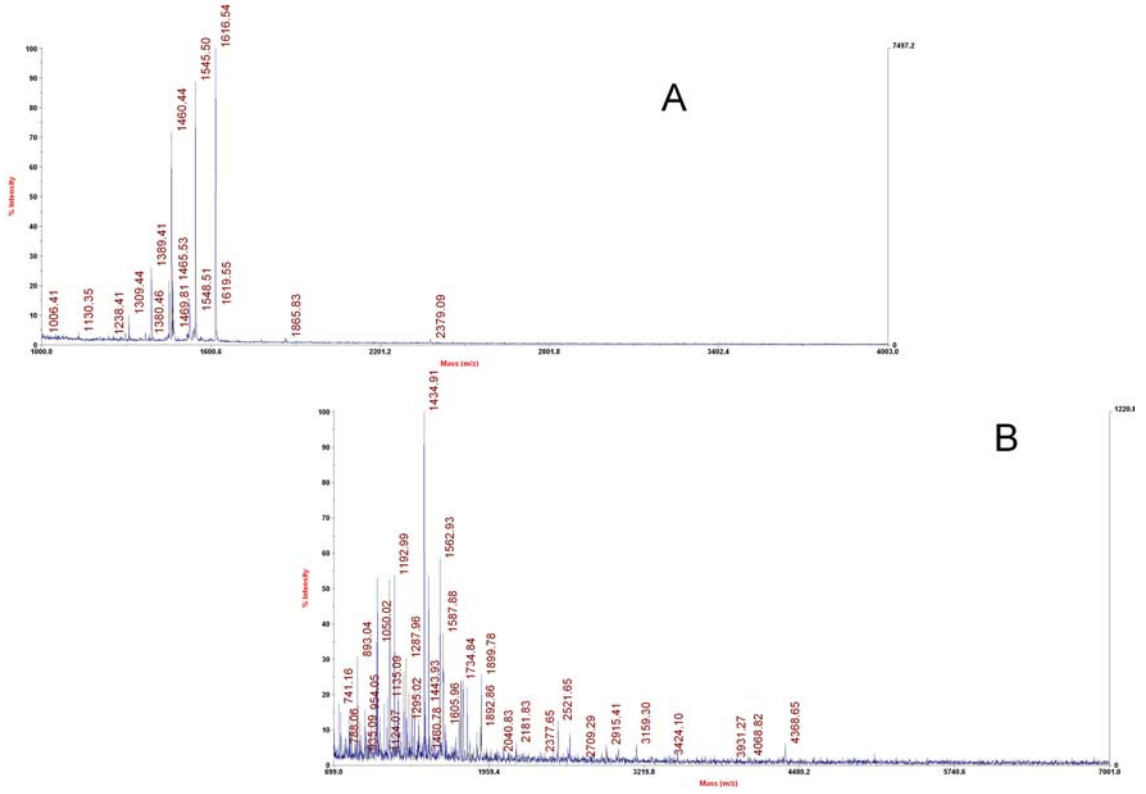


Figure. 15 A-B MALDI-MS analysis of enriched peptide mixture from serum and saliva respectively.

The selective analysis of phosphopeptides by MS-MS was then applied. MS-MS selective analysis were reiterated for two times for two different donators. High selectivity of precursor ion scan was emphasized for the analysis of free phosphopeptides of biological fluids.

In these experiments a higher number of phosphopeptides with respect to those detected by using neutral loss analysis was identified despite inferior MS-MS quality spectra. This is probably due to the worse stability of the spray switching the polarity, the length of the total acquisition cycle and the different charged state of the ion parent between precursor ion and neutral loss analysis. In this case the neutral loss scan was performed at 49 Da corresponding to phosphoric acid neutral loss from doubly charged peptides whereas the precursor ion scan was performed on doubly and triply charged peptides.

Figure 16 panel A shows the resulting survey precursor ion spectrum in negative ion mode related to 35 minutes of one of the serum free phosphopeptides, the resulting ER scan performed in the positive ion mode and the corresponding EPI scan performed on the precursor triply charged ion at m/z 515.86. Figure 16 panel B shows the resulting survey neutral loss of 49 Da at the same retention time of the same peptide but doubly charged; the resulting ER scan performed in the positive ion mode and the corresponding EPI scan performed on the precursor at m/z 773.31.

The specificity of the precursor ion and neutral loss scans is clearly indicated by the clarity of the survey MS spectrum.

The comparison between the MS-MS spectra shows that the EPI spectra arising from doubly charged peptide is richer in signals and better quality with

respect to the same fragment spectrum than the triply charged peptide. Figure 16 shows the partial sequence of this peptide.

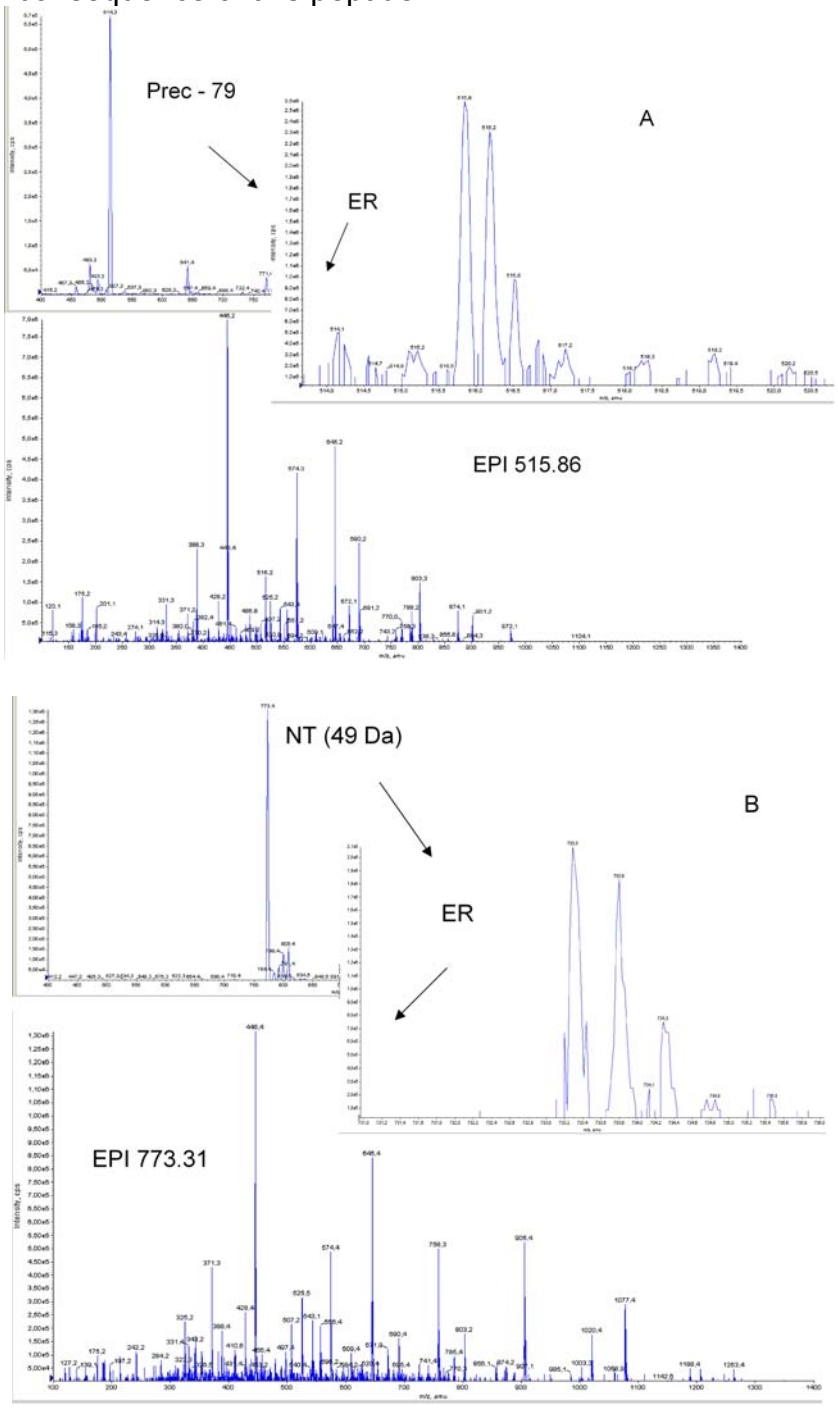


Figure. 15 A-B Comparison between Precursor ion scan and Neutral loss analysis of the same peptide with sequence DSpGEGDFLAEGGGVR respectively.

As a whole, neutral loss analysis essentially confirmed previous results but allowed us to identify some phosphopeptides not detected before. However a minor number of analytes was detected by neutral loss scan compared to precursor ion scan analysis probably due to the presence of not phosphorylated peptides that unspecifically interacted with IMAC resin and gave rise to a loss of 49 Da during neutral loss analysis. All serum results are summarized in table 8.

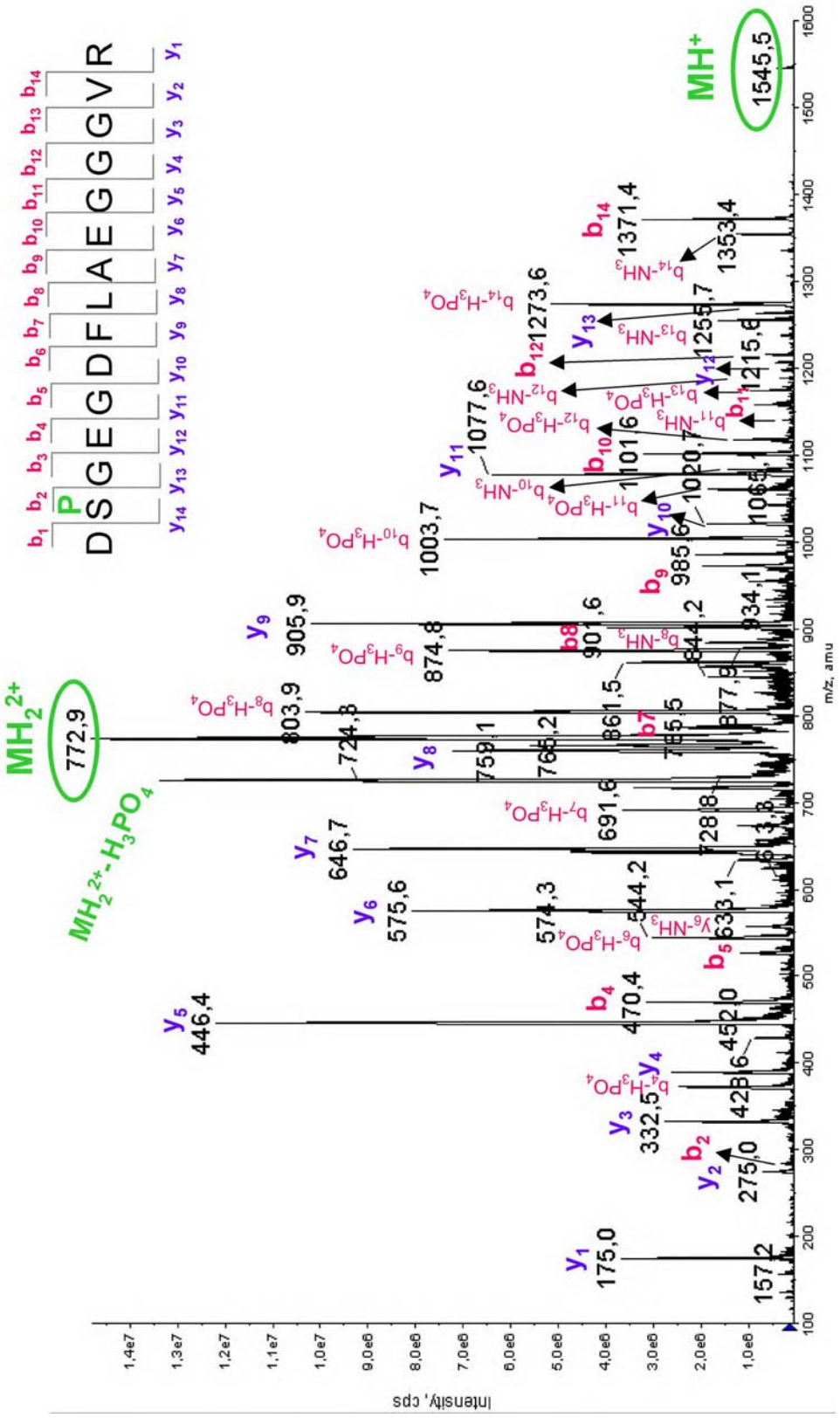


Figure. 16 MS-MS spectrum of the signal double charged at m/z 773.33 with sequence DSpGEGDFLAEGGGVR.

Table 8 Identified free phosphopeptides from serum. Reported phosphorylation site in literature were underlined in yellow.

Sequence peptide	Protein	Modification
DSGEGDFLAEGGGV	P02671 Fibrinogen α-chain	pS
ADSGEGDFLAEGGGV	P02671 Fibrinogen α-chain	pS
ADSGEGDFLAEGGGVR	P02671 Fibrinogen α-chain	pS
DSGEGDFLAEGGGVR	P02671 Fibrinogen α-chain	pS

When the experiments were carried out by using an enriched saliva sample, many free phosphopeptides belonging to salivary proteins were observed in neutral loss and precursor ion scan analysis.

As an example, the fragment identified at m/z 538.24 corresponding to statherin precursor was shown in Figure 17. In addition to a dominant H3PO4 loss (m/z 98/2) ion at m/z 489.4 characteristic of a doubly charged monophosphorylated peptide, a continuous series ranging from y1 to y6 ions were observed. Moreover, a series of b ions, namely, b3-H3PO4 , b4-H3PO4, and y7-H3PO4, which resulted from fragmentation of the β-eliminated 489.4 m/z ion were detected. In this case the loss of phosphoric acid was complete and assignment of correct phosphorylation site was hard.

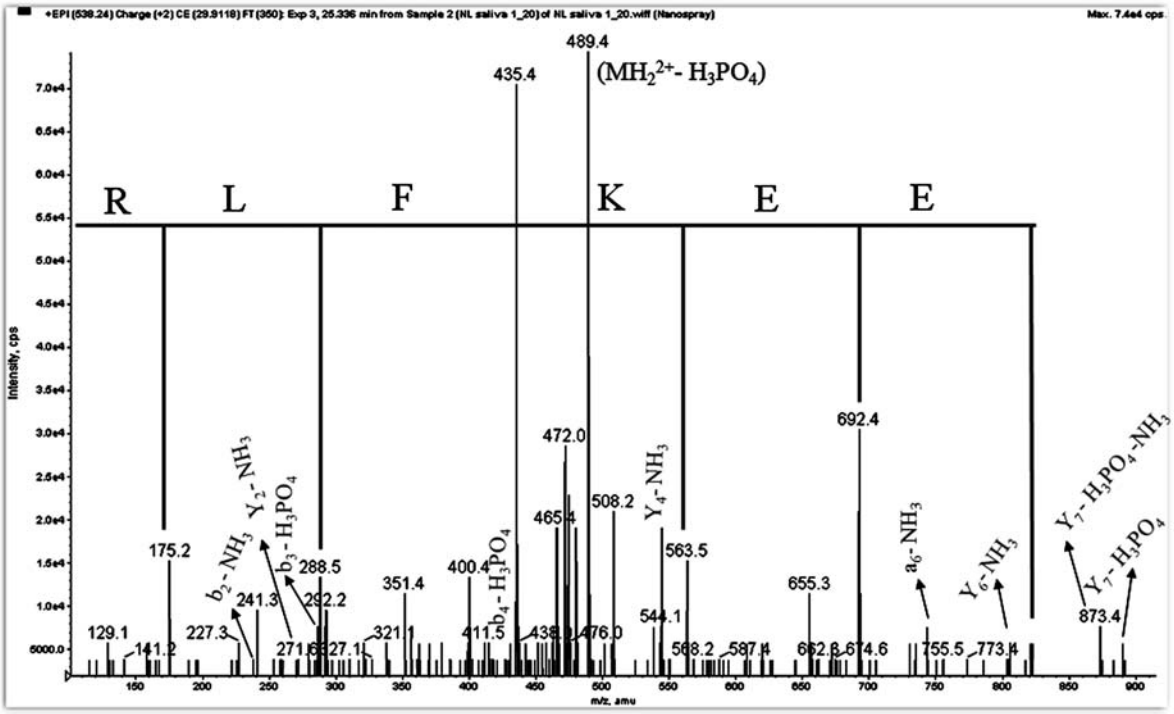


Figure 17. MS-MS spectrum of a free salivary monophosphorylated peptide at m/z 538.24

Moreover, the MS analysis led to the detection of another free phosphopeptide identified as DpSpSEEKFLR within statherin sequence whose spectrum was shown in figure 18. As indicated, the spectrum exhibited two diagnostic ions for a doubly charged diphosphopeptide: a dominant H_3PO_4 loss (m/z 98/2) ion at m/z 586.8 and a second H_3PO_4 loss (m/z 98/2) at m/z 537.8. A continuous series ranging from y7 to y10 ion consistent with a phosphoserine residue at both positions 2 and 3 were observed. It is clear that the interpretation of this spectrum is complicated by presence of different fragmentation series from

MH_2^{2+}, MH_2^{2+}- H_3PO_4 and MH_2^{2+}- $2H_3PO_4$ often unexplained by conventional b- and y- type ions.

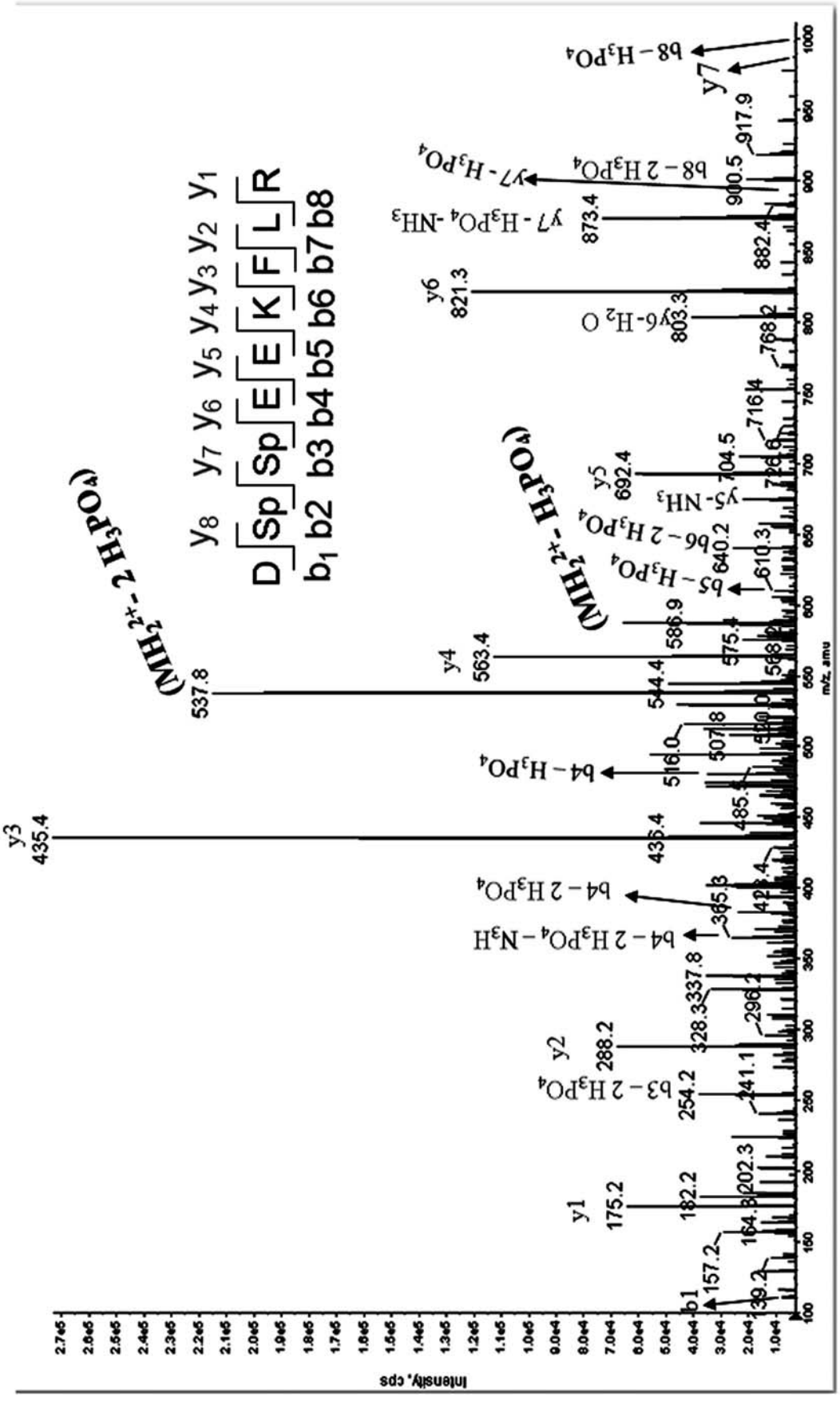


Figure 18. MS-MS spectrum of a free salivary diphosphorylated peptide at m/z 635.74.

Whereas, in other cases the loss of phosphoric acid was complete thus producing a spectrum easier to interpret but the challenging in assignment of the correct phosphorylation site still remained. As example, figure 19 shows the MS-MS spectrum of a double charged phosphopeptide at m/z 731.27, whose partial sequence was GGDpSQFIDEER identified by MASCOT analysis, with a complete loss of phosphoric acid and with an unique serine residue within sequence.

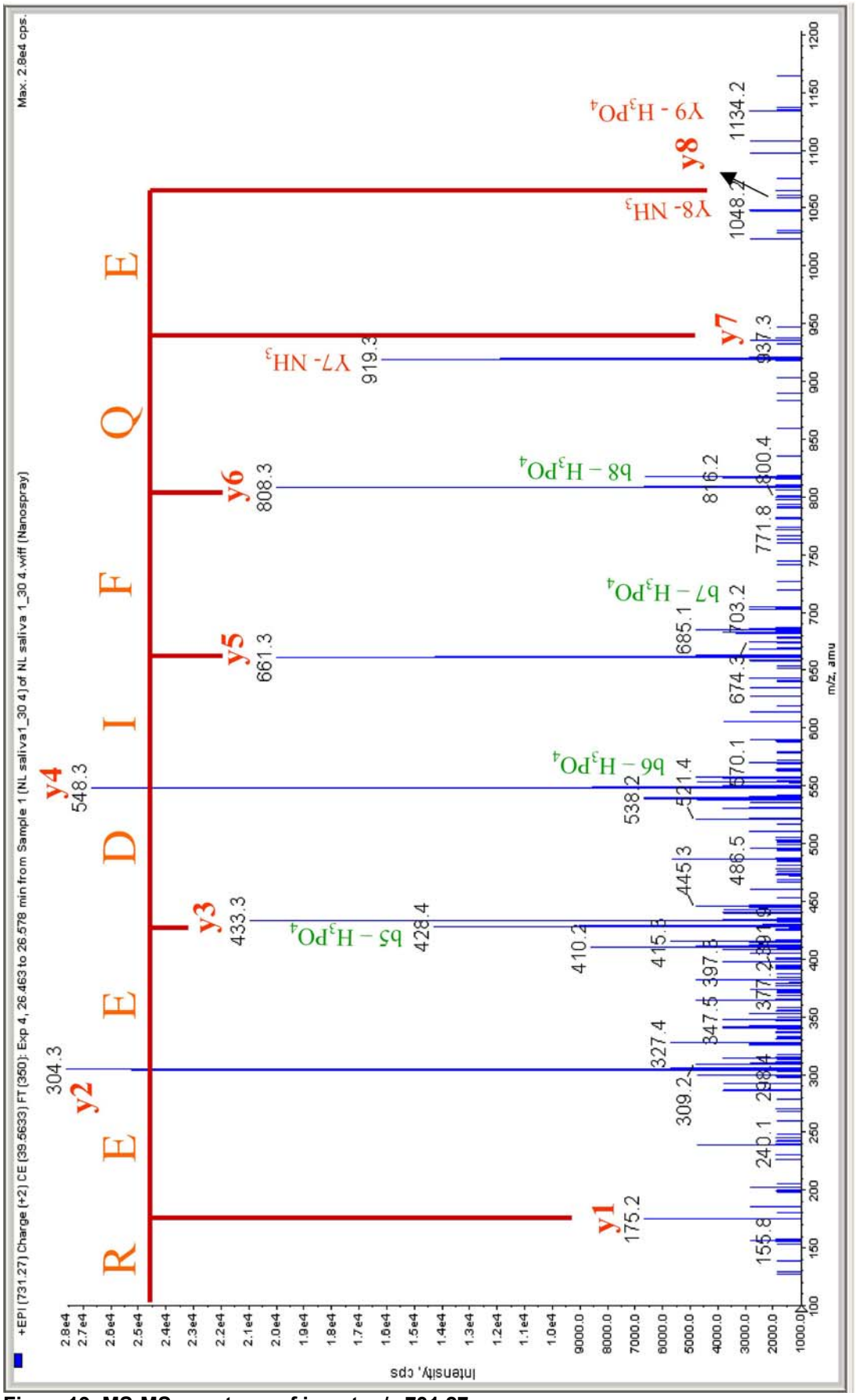


Figure19. MS-MS spectrum of ion at m/z 731.27.

A total of 9 free phosphopeptides and six phosphorylation sites within the four protein sequences were identified (tab.11). In particular, free phosphopeptides from salivary acidic praline-rich phosphoprotein ½ precursor (PRP), basic salivary praline-rich protein 2, statherin precursor and histatin were detected.

Table 11. Identified free phosphopeptides from saliva. Reported phosphorylation sites in literature were underlined in yellow. New phosphorylation sites were underlined in red.

Sequence	Protein	Modification
GGDSEQFIDEER	P02810 Salivary acidic proline-rich phosphoprotein 1/2 precursor (PRP-1/PRP-2)	1P
GGDSEQFIDEERQ	P02810 Salivary acidic proline-rich phosphoprotein 1/2 precursor (PRP-1/PRP-2)	1P
PGKPQGPPPQGGSKSRSARS	P02812 Basic salivary proline-rich protein 2	2P
PGKPQGPPPQGGSKSRSARS	P02812 Basic salivary proline-rich protein 2	2P
SSEEKFLR	P02808 Statherin precursor	1P
SSEEKFLR	P02808 Statherin precursor	2P
DSSEEKFLR	P02808 Statherin precursor	1P
DSSEEKFLR	P02808 Statherin precursor	2P
SHEKRHHGYR	P15515 Histatin-1 precursor	1P

It is worth that the analysis of free peptides is still challenging because of the missing knowledge of the C-terminal aminoacid. The fragmentation of tryptic peptides is optimised thanks to the presence of a basic aminoacid at C-terminal residue. Unfortunately many free peptides did not present a lysine or arginine at C-terminal residue. In these analysis many peptides were found having different cleavage sites and this was reflected in the low quality of fragmentation spectra.

No phosphopeptide was identified in urine. The analyzed urine samples belonged to healthy donors. The porous wall of the glomerular capillaries in the kidney normally allows only proteins with a molecular weight < 65 kDa into the tubular system, whereas larger proteins stay in the blood stream. The tubular cells subsequently reabsorb the smaller proteins, and only low concentrations of proteins are normally present in the normal urine (39).

In conclusion, the combination of the IMAC improved protocol with selective “precursor ion” and constant “neutral loss” triple quadrupole scan modes confers the high sensitivity allowing rapid peptide identification and characterization at low concentrations.

The unique capabilities of the Q-q-Q linear ion trap in selective ion scans are due to the decoupling and spatial separation of the ion isolation, fragmentation, and mass analysis steps of the instrument.

III.7 Conclusions

Tumor markers are critical in the diagnosis of diseases and monitoring of patients.

Powerful analytical technologies and, especially, mass spectrometry, have allowed identification of new markers. Recently, it has been found that not only proteins but also peptides and their modifications can be indicators of early pathogenic processes. The extraordinary power of mass spectrometry in identifying and quantifying peptides in complex biological mixtures offers opportunities for developing novel technologies for diagnosis of cancer and other diseases (35).

Since biological fluids contain high concentration of large abundant proteins such as albumin, uromodulin, histatin, it is generally accepted that these proteins must be

removed before analysis to allow comparison of other less abundant proteins. Moreover, enrichment methods are required to detect phosphospecies.

In this study IMAC chromatography and selective mass spectrometry scans were improved to study phosphopeptidomics in biological fluids. One strategy to increase the sensitivity and selectivity was to use a linear ion trap to analyze the purified phosphopeptides.

The application of this IMAC protocol coupled with MS-MS experiments significantly improved the analysis of phosphopeptides in biological fluids. Many peptides were detected in the enriched extract samples when submitted to the integrated strategy, whereas they resulted absent in the initial extract samples.

In 2003, Marshall et al. claimed that peptides from the sera of normal individuals and patients who suffered myocardial infarction can produce MALDI-TOF patterns that provide an accurate diagnosis of myocardial infarction (23). In some patients who had occluded coronary arteries phosphorylated fibrinogen Aα levels increased as compared with patients who had patent coronary arteries. These results provide some support for the idea that an increased synthesis of fibrinogen in circulation may result in a procoagulant tendency. Thus, phosphorylated fibrinogen Aα content may serve as a risk index for thrombosis (24).

Human saliva contains a large number of proteins which have important physiological functions. Nowadays, there is a growing interest for these proteins, due their potential usefulness for a proteomic analysis in the diagnosis of many different diseases. The knowledge of the phosphopeptides can give rise an important information about biochemical processes revealing potential target for therapy of diseases.

Saliva is usually supersaturated with respect to the basic calcium phosphate salts that form dental enamel, a property which provides protection for the teeth. An undesirable side effect of this is that unwanted precipitation of solid calcium phosphate may occur in the salivary glands, the fluid of the mouth, or on dental enamel (25). However, this is avoided by the presence of proteins such as statherin (26), a phosphoprotein that prevents calcium phosphate precipitation and growth of hydroxyapatite crystals (25). The latter activity is shared by acidic proline-rich proteins (APRPs)1 (27). Statherin and APRPs loose these activities upon dephosphorylation (25). Moreover, the ability of APRPs to bind calcium and thereby aid in maintaining the concentration of ionic calcium in saliva also depends on phosphorylation of the proteins (28-29). In addition, dephosphorylation of APRPs decreases their binding to hydroxyapatite (30) thus preventing them from taking part in formation of dental pellicle, the thin layer of adsorbed proteins found on teeth which may provide protection against demineralization (31). Moreover, APRPs may be relevant modulators of bacterial colonization of the teeth (32- 33-34).

In conclusion, the results obtained in this work demonstrated that the improved Immobilized Metal Affinity Chromatography and Selective tandem mass spectrometry of pre-fractionated biological fluids allowed the identification of potential biomarkers that may lead to the development of tests for screening of high-risk individuals. This method should be widely applicable to the searching for biomarker in clinical diagnosis and therapy.

III.8 References

1. Porath J, Carlsson J, Olsson I, Belfrage G. Metal chelate affinity chromatography, a new approach to protein fractionation. Nature. 1975. 18;258(5536):598-9.
2. Shihua L. and Chhabil D. Iron(III)-Immobilized Metal Ion Affinity Chromatography and Mass Spectrometry for the Purification and Characterization of Synthetic Phosphopeptides. Anal. Biochem. 1999. 270: 9–14.
3. Ueda EK, Gout PW, Morganti L. Current and prospective applications of metal ion-protein binding. J Chromatogr A 2003. 21;988(1):1-23.
4. M.O. Collins, L.Yu., H.Husi, W.P. Blackstock, J.S. Choudhary, S.G.N. Grant. Robust Enrichment of Phosphorylated Species in Complex Mixtures by Sequential Protein and Peptide Metal-Affinity Chromatography and Analysis by Tandem Mass Spectrometry. Sci. STKE. 2005. 298: 16.
5. Steen H, Kuster B, Mann M. Quadrupole time-of-flight versus triple-quadrupole mass spectrometry for the determination of phosphopeptides by precursor ion scanning. J Mass Spectrom. 2001. 36(7):782-90.
6. Steen H, Kuster B, Fernandez M, Pandey A, Mann M. Detection of tyrosine phosphorylated peptides by precursor ion scanning quadrupole TOF mass spectrometry in positive ion mode. Anal Chem. 2001. 1;73(7):1440-8.
7. Carr SA, Huddleston MJ, Annan RS. Selective detection and sequencing of phosphopeptides at the femtomole level by mass spectrometry. Anal Biochem. 1996. 1;239(2):180-92.
8. Bruins AP, Covey TR, Henion JD. Ion spray interface for combined liquid chromatography/atmospheric pressure ionization mass spectrometry. Anal Chem. 1987. 59:2642–2646.
9. Schulte, I.; Tammen, H.; Selle, H.; Schulz-Knappe, P. Peptides in body fluids and tissues as markers of disease. Expert Rev. Mol.Diagn. 2005. 5: 145-157.
10. Schrader, M.; Schulz-Knappe, P. Peptidomics technologies for human body fluids. Trends Biotechnol. 2001. 19: S55-60.
11. Jimenez-Reyes M, Sanchez-Aguirre FJ. Sodium and chlorine concentrations in mixed saliva of healthy and cystic fibrosis children. Appl Radiat Isot. 1996. 47(3):273-7.
12. Ebert, M. P.; Meuer, J.; Wiemer, J. C.; Schulz, H. U.; Reymond, M. A.; Traugott, U.; Malfertheiner, P.; Rocken, C. J. Proteome Res. 2004. 3: 1261-1266.
13. Li, J.; Zhang, Z.; Rosenzweig, J.; Wang, Y. Y.; Chan, D. W. Clin. Chem. 2002. 48: 1296-1304.
14. Gruhler A, Olsen JV, Mohammed S, Mortensen P, Faergeman NJ, Mann M, Jensen ON. Mol Cell Proteomics. Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. 2005. 4(3):310-27.
15. Giorgianni F, Beranova-Giorgianni S, Desiderio DM. Identification and characterization of phosphorylated proteins in the human pituitary. Proteomics. 2004. 4(3):587-98.
16. Ballif BA, Villen J, Beausoleil SA, Schwartz D, Gygi SP. Phosphoproteomic analysis of the developing mouse brain. Mol Cell Proteomics. 2004. 3(11):1093-101.
17. Kjellstrom S, Jensen ON. Phosphoric acid as a matrix additive for MALDI MS analysis of phosphopeptides and phosphoproteins. Anal Chem. 2004 1;76(17):5109-17.
18. Zappacosta, F., Huddleston, M.J., Karcher, R.L., Gelfand, V.I., Carr, S.A. and Annan,R.S. Improved sensitivity for phosphopeptide mapping using capillary column

HPLC and microionspray mass spectrometry: comparative phosphorylation site mapping from gel-derived proteins. Anal. Chem. 2002. 74: 3221-3231.
19. Steen, H., Kuster, B., Fernandez, M., Pandey, A. and Mann, M. Detection of tyrosine phosphorylated peptides by precursor ion scanning quadrupole TOF mass spectrometry in positive ion mode Anal. Chem. 2001 73: 1440-1448.
20. LeBlanc, J.C.Y., Hager, J.W., Ilisiu, A.M.P., Hunter, C., Zhong, F., Chu,I. Unique scanning capabilities of a new hybrid linear ion trap mass spectrometer (Q TRAP) used for high sensitivity proteomics applications Proteomics. 2003. 3: 859-869.
21. Shou W, Verma R, Annan RS. Mapping phosphorylation sites in proteins by mass spectrometry. Methods Enzymol. 2002. 351:279–296.
22. Chertov O, Biragyn A, Kwak LW, Simpson JT, Boronina T, Hoang VM, Prieto DA, Conrads TP, Veenstra TD, Fisher RJ. Organic solvent extraction of proteins and peptides from serum as an effective sample preparation for detection and identification of biomarkers by mass spectrometry. Proteomics. 2004. 4(4):1195-203.
23. Marshall, J.; Kupchak, P.; Zhu, W.; Yantha, J.; Vrees, T.; Furesz, S. Processing of serum proteins underlies the mass spectral fingerprinting of myocardial infarction. J. Proteome Res. 2003. 2: 361-372.
24. Maurer MC, Peng JL, An SS, Trosset JY, Henschen-Edman A, Scheraga HA. Baker. Structural examination of the influence of phosphorylation on the binding of fibrinopeptide A to bovine thrombin. Biochemistry. 1998. 28:37(17): 5888-902.
25. Hay, D. I., Moreno, E. C., and Schlesinger, D. H. Inorg. Perspec. Biol. Med. 1979. 2: 271-285.
26. Schlesinger DH, Hay DI. Complete covalent structure of statherin, a tyrosine-rich acidic peptide which inhibits calcium phosphate precipitation from human parotid saliva. J Biol Chem. 1977. 10;252(5):1689-95.
27. Naurato N, Wong P, Lu Y, Wroblewski K, Bennick A. Interaction of tannin with human salivary histatins. J Agric Food Chem. 1999. 47(6):2229-34.
28. Chan M, Bennick A. Proteolytic processing of a human salivary proline-rich protein precursor by proprotein convertases. Eur J Biochem. 2001. 268(12):3423-31.
29. Hay DI, Carlson ER, Schluckebier SK, Moreno EC, Schlesinger DH. Inhibition of calcium phosphate precipitation by human salivary acidic proline-rich proteins: structure-activity relationships. Calcif Tissue Int. 1987. 40(3):126-32.
30. Bennick A, Cannon M, Madapallimattam G. Factors affecting the adsorption of salivary acidic proline-rich proteins to hydroxyapatite. Caries Res. 1981;15(1):9-20.
31. Moreno EC, Zahradnik RT. Demineralization and remineralization of dental enamel. J Dent Res. 1979. 58(Spec Issue B):896-903.
32. Gibbons RJ, Hay DI, Cisar JO, Clark WB. Adsorbed salivary proline-rich protein 1 and statherin: receptors for type 1 fimbriae of Actinomyces viscosus T14V-J1 on apatitic surfaces. Infect Immun. 1988. 56(11):2990-3.
33. Gibbons RJ, Hay DI, Schlesinger DH. Delineation of a segment of adsorbed salivary acidic proline-rich proteins which promotes adhesion of Streptococcus gordonii to apatitic surfaces. Infect Immun. 1991. 59(9):2948-54.
34. Bennick A, Cannon M, Madapallimattam G. The nature of the hydroxyapatite-binding site in salivary acidic proline-rich proteins. Biochem J. 1979. 1;183(1):115-26.
35. Hyoung-Joo Lee, Eun-Young Lee, Min-Seok Kwon and Young-Ki Paik Biomarker discovery from the plasma proteome using multidimensional fractionation proteomics. Current Opinion in Chemical Biology 2006, 10:42–49.

36. Francesco Giorgianni; Sarka Beranova-Giorgianni; Dominic M. Desiderio. Identification and characterization of phosphorylated proteins in the human pituitary Proteomics 2004, 4, 587–598.
37. Ficarro SB, McCleland ML, Stukenberg PT, Burke DJ, Ross MM, Shabanovwitz J, Hunt DF,White FM. Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces Cerevisiae. Nat Biotechnol. 2002. 3:301-5.
38. D'Angelo A, Garzia L, Andre A, Carotenuto P, Aglio V, Guardiola O, Arrigoni G, Cossu A, Palmieri G, Aravind L, Zollo M. Prune cAMP phosphodiesterase binds nm23-H1 and promotes cancer metastasis. Cancer Cell. 2004. 5:137-49.
39. Ottiger C, Savoca R, Yurtsever H, Huber AR. Increased sensitivity in detecting renal impairments by quantitative measurement of marker protein excretion compared to detection of pathological particles in urine sediment analysis. Clin Chem Lab Med. 2006. 44(11):1347-1354.

IV. Selective detection and identification of phosphopeptides by dansyl MS/MS/MS fragmentation

IV.1 Dansyl Chemistry

The challenge of large-scale functional genomics projects is to build a comprehensive map of the cell. In recent years, mass spectrometry (MS) has been recognized as the best tool for proteomic analysis.
Several group have reported the analytical potential of chemical derivatization in conjunction with mass spectrometry for the analysis of synthetic and biological compounds. Some derivatizing agent improves the ionization and the sensitivity of analyses by MS.

Dansyl chloride (DANS-Cl) is a fluorescent derivatizing agent commonly used in the analysis of proteins and other primary and secondary amines. The presence of a dimethylamino group in this compound makes it amenable to protonation under acidic conditions. Dansylation of compounds that are not readily ionized results in higher sensitivity in MS analysis.

Since 1968 (1) the group, where I have sustained my phD work, discovered some interesting properties of 1-dimethylaminoaphthalene-5-sulphonyl- (DANS) amino acid derivatives. Since DANS derivatives give rise to a very intense fragment at m/z: 170, the “metastable refocusing” (a primordial precursor ion scanning) first was applied for the analysis of amine mixtures (2-3).

Recently, revisiting these previous findings by using soft ionization procedures (ESI and MALDI), we have developed new strategies based on the use of dansyl chloride and linear ion trap aimed at identification of bacterial surface-exposed proteins and post- translational modification (PTMs) both in qualitative and quantitative terms.

IV.2 Aim of the project

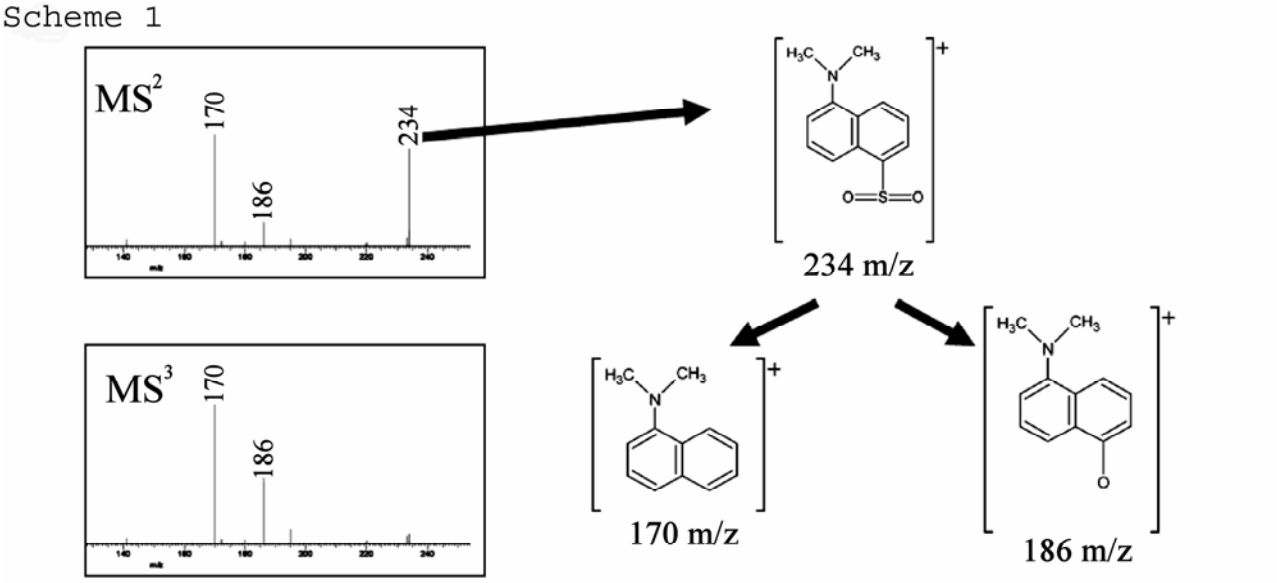
Selective detection of phosphopeptides from proteolytic digests is a challenging and highly relevant task in many proteomics applications. Because of the impact of enzyme-catalyzed phosphorylation and dephosphorylation on the function of many proteins (4), there is a demand to improve the methods for a reliable analysis of phosphopeptides. Phosphoproteins often exist in low abundance in cells besides to corresponding non-phosphorylated parent forms, and they need selective isolation or enrichment before identification. Moreover, phosphopeptides are easily suppressed in mass spectrometry in both MALDI and ESI ionization modes. Furthermore, the low stoichiometry of phosphorylation in proteins makes important to increase the selectivity in the analysis (5).

Treatment with an appropriate derivatizing agent has the potential to improve the ionization of a compound and therefore the detectability of analytes by MS. Here it was suggested that derivatization may include chemical moiety capable to give rise to characteristic fragment ions, thus allowing the exploitation of the enormous potential of MS^n analysis (6-7). For the first time the novel acronym RIGhT (Reporter Ion Generating Tag) was introduced to indicate the general derivatization method

based on dansyl labelling of the targeted residues. In the case of phosphopeptide the procedure consists in the β-elimination of phospho-Ser/-Thr residues and the Michael addiction of the resulting α,β-unsatured residues with DAN-SH. The dansyl derivatization introduces: a) a basic secondary nitrogen into the molecule that enhances the efficiency of signal ionization; b) a dansyl moiety that fragments according to previous data (2-8). Using the great capabilities of a new hybrid mass spectrometer equipped with a linear ion trap analyzer, one can take advantage of the distinctive m/z 170 and m/z 234 fragments in MS^2 and the diagnostic 234→170 fragmentation in MS^3 mode (scheme1).

As a whole this approach, first applied on the identification of phosphorylation sites in model phosphopeptides, promise to enrich greatly our global view of the dynamic changes occurring even in complex matrices such as human serum. It is worth considering that O-glycosylated residues also can be targeted with the DAN reagent thus widening the range of effectiveness of the method.

Scheme 1



Scheme 1. Characteristic DANSIL-fragmentation.

IV.3 Materials and Methods

Proteomic grade trypsin, α-casein, myoglobin, bovine serum albumin, insulin, trasferrin, ribonuclease A, carbonic anhydrase, lysozime, glutathione S-transferase, enolase and tributylphosphine solution (stock solution 200mM in glass vial, sealed under Ar) were purchased from Sigma (St. Louis, MO, USA). Dansyl chloride and (NH2CH2CH2S)2 (cystamine) was of the highest purity available from SIGMA (BioChemika >98.0 %). Trifluoroacetic acid HPLC grade was from Carlo Erba (Milano, Italy). All used solvents were of the highest purity available from Baker (Phillipsburg, NJ). All other reagents and proteins were of the highest purity available from Sigma.

Phosphate group modification

The β-eliminated peptide mixture (see Material and methods section of capitol 2) was modified using DAN-SH via Michael-type addition. Briefly, DANSS was obtained by reaction of Dansil chloride (0.1 mg/mL dissolved in acetonitrile) with cystamine (molar ratio 3:1). Reaction was carried at 60° C for 2 hr. The product was purified by RP-HPLC using an Agilent Zorbax C8 column (4,6mm X 150mm) (Palo Alto, California) using a 5% to 65% linear gradient in 30 min from water to acetonitrile (ACN) and verified by ESI-MS-MS analysis. Dried DANS-S was dissolved in 10 mM Tris/HCl Buffer (~15 nmol/μL, pH 8.5) and reduced in presence of 20 mM tributylphosphine. Reaction was carried out for 30 min at room temperature under nitrogen. Product was purified by RP-HPLC in the same condition previously described and verified by ESI/MSMS analysis. 100 μL of DANSH solution (0.1 mg/mL, dissolved in ACN/Water 3:1) was then added to an equal volume of the β-eliminated peptide mixture solution (1 nmol/mL, in deionized water) in presence of 1 mM tributylphosphine. Tris 2 M was added to the sample at a final concentration of 100 mM (pH ~11). The reaction proceeded under nitrogen for 18 hours at 50 °C

MALDI-TOF mass spectrometry

MALDI-TOF mass spectra were recorded using a Voyager DE-PRO mass spectrometer (Applied Biosystem, Framingham, USA). Prior of analysis, peptide mixtures were purified using ZipTip pipette from Millipore (Billerica, MA, USA) using the recommended purification procedure. The peptides were eluted using 20 μL of 50% ACN, 0.1% TFA in water. A mixture of the eluted peptide solution and α-cyano-hydroxycinnamic acid (10 mg/mL in ACN: 0.1% TFA in water, 2:1, v/v) was applied to the metallic sample plate and dried at room temperature. Mass calibration was performed using a mixture of peptides from Applied Biosystem, containing des-Arg1-Bradykinin, Angiotensin I, Glu1-Fibrinopeptide B, ACTH (1-17), ACTH (18-39) and Insulin (bovine) as external standards. Raw data were analysed using Data Explorer software provided by Applied Biosystem and reported as monoisotopic masses.

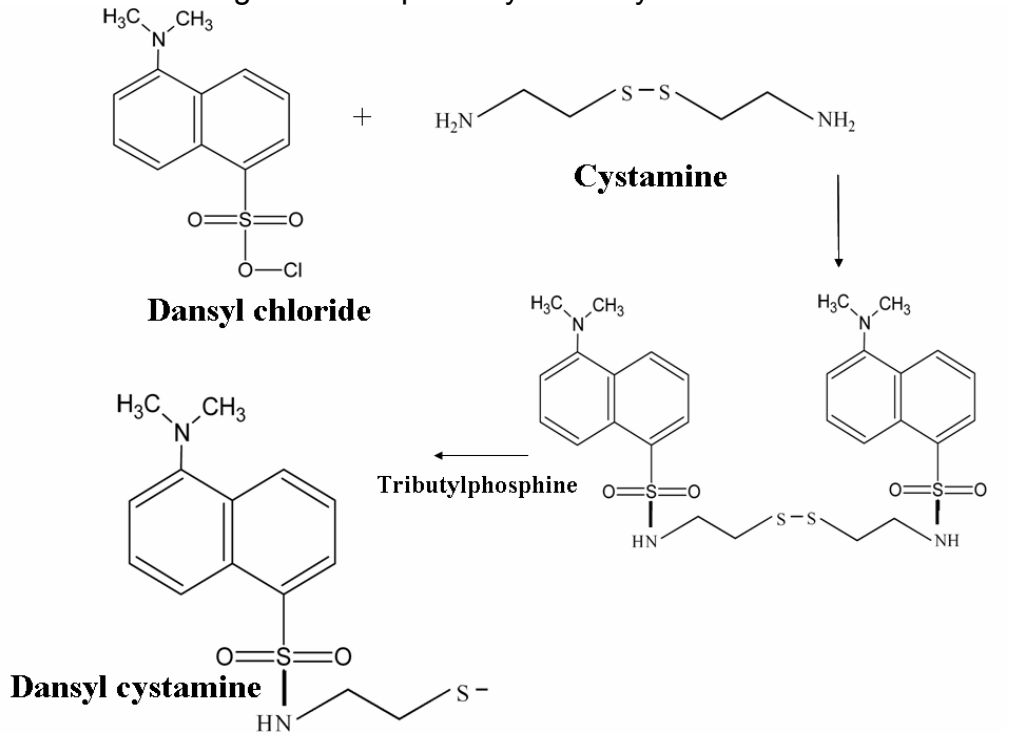
nanoLC Mass Spectrometry

Peptide mixture, obtained as previously described, was analysed by LCMS analysis using a 4000Q-Trap (Applied Biosystems) equipped with a linear ion trap (Scheme 1) coupled to an 1100 nano HPLC system (Agilent Technologies). The mixture was loaded on an Agilent reverse-phase pre-column cartridge (Zorbax 300 SB-C18, 5x0.3 mm, 5 μm) at 10 μL/min (A solvent 0.1% formic acid, loading time 7 min). Peptides were separated on a Agilent reverse-phase column (Zorbax 300 SB-C18, 150 mm X 75μm, 3.5 μm), at a flow rate of 0.2 μL/min with a 5 to 65% linear gradient in 60 min (A solvent 0.1% formic acid, 2% ACN in water; B solvent 0.1% formic acid, 2% water in ACN). Micro ionspray source was used at 2.4 kV with liquid coupling, with a declustering potential of 50 V, using an uncoated silica tip from NewObjectives (Ringoes, NJ) (O.D. 150 μm, I.D. 20 μm, Tip. Diameter. 10 μm). Spectra acquisition was based on a survey Precursor Ion Scan. for the ion m/z.170. The Q1 quadrupole was scanned from m/z 500 to m/z 1000 in 2 sec, and ions were fragmented in q2 using a linear gradient of collision potential from 30 to 70 V. Finally, Q3 is settled to transmit only ions at m/z 170. This scan mode was followed by an enhanced resolution experiment (ER) for the ions of interest and then by MS3 and

MS^2 acquisitions of the two most intense ions. MS^2 spectra were acquired using the best collision energy calculated on the bases of m/z values and charge state (rolling collision energy). MS^3 spectra were acquired using Q0 Trapping, with a trapping time of 150 ms and an activation time of 100 ms, scanning from m/z 160 to 240. The Q0 Trapping tool is able to increase the sensitivity of the scan in the different ion trap scan types used. Data were acquired and processed using Analyst software (Applied Biosystems).

IV.4 Results and discussion

In an attempt to improve and simplify the plethora of methods so far suggested (10) a new approach was set up to label selectively phospho-Ser/-Thr residues by exploiting the features of novel linear ion trap mass spectrometer (11). Dansyl chloride is known to react with ε-amino group of lysine as well as the N-terminal amino group (9). Here, using chemical manipulations coupled with dansyl chloride labelling, a methodology capable of large-scale proteomic profiling of phosphorylation sites was shown. This method is based a) on the selective modification of pSer/Thr residues with a novel dansyl derivative of $NH_2(CH_2)_2SH$, cysteamine, prepared according to material and methods sections and shown in Scheme 2 and b) on the selective detection and identification of labelled peptides by exploiting the characteristic fragmentation pathway of dansyl-derivatives.



Scheme 2. Synthesis of DANS-SH.

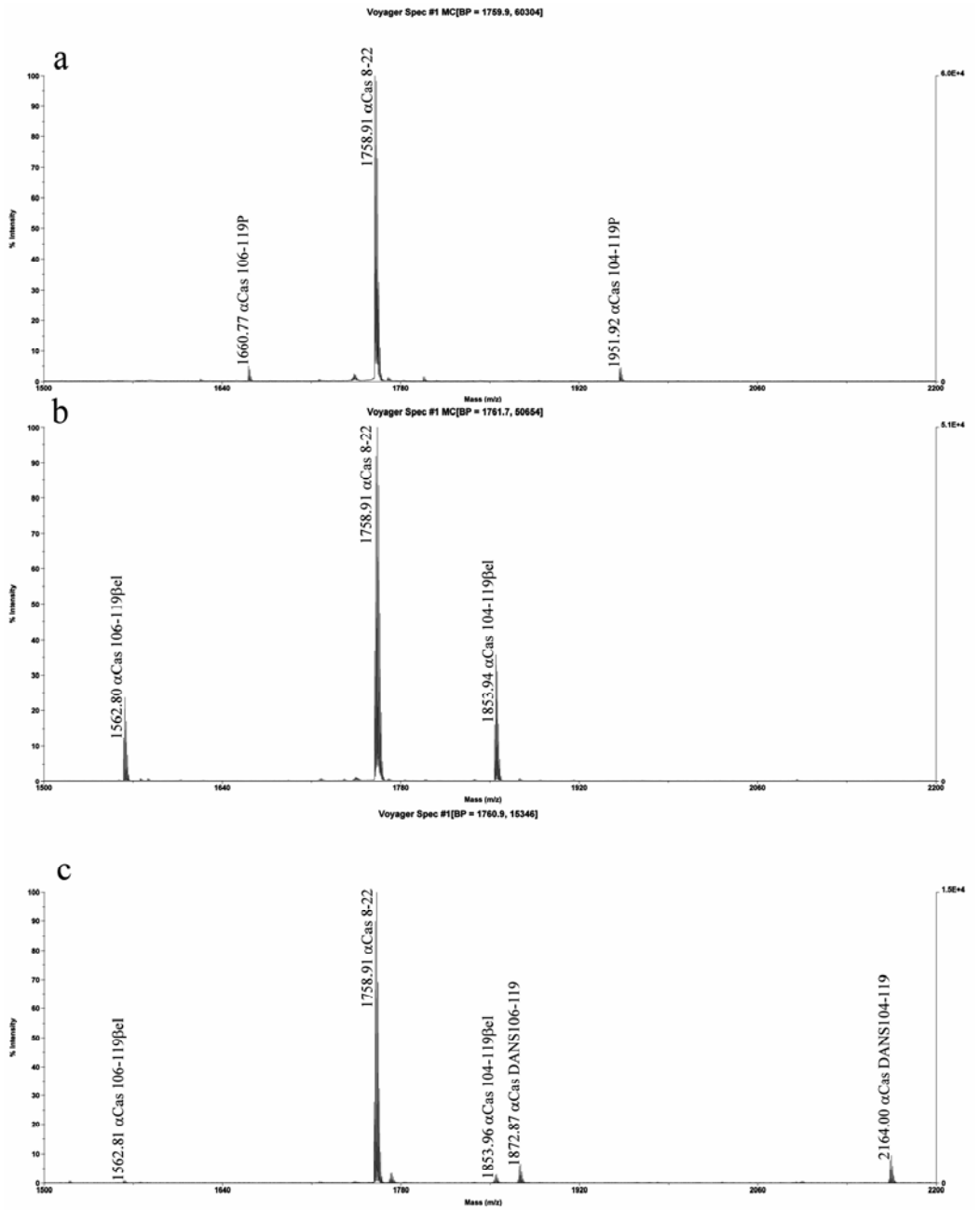


Figure 1. a: Partial MALDIMS spectrum of the trypsin digest of α--casein. b: Partial MALDIMS spectrum of the β-eliminated peptide mixture. The mass signals related to the phosphorylated peptides are shifted of 98 Da due to the β-elimination reaction. c: Partial MALDIMS spectrum of the β-eliminated peptide mixture after N-(2-mercaptoethyl)dansylamide addition. The signals corresponding to the β-eliminated peptides show a mass increment of 310 Da.

This approach was first applied to α-casein as model phosphoprotein. Thus, the phosphate moieties from a tryptic α-casein digest were removed from pSer and pThr via barium hydroxide ion-mediated β-elimination as already reported (11). The peptide mixture was then analysed by MALDIMS, using a Voyager DE-PRO mass spectrometer. Prior of analysis, peptide mixtures were purified using ZipTip pipette, using the recommended purification procedure.

The MALDI-MS analysis showed the occurrence of signals corresponding to the β-eliminated peptides. As an example, the ion at m/z 1854.3 was assigned to the peptide 104-119 of α-casein occurring 98 Da lower than that expected, indicating that

the pSer115 was converted in dehydroalanyl after β-elimination reaction (Figure 1). The α,β-unsatured residues are Michael acceptor, which can readily react with a nucleophile. The β-eliminated peptide mixture was modified with DANSH via Michael-type addition.

The yield of addition reaction was monitored via MALDIMS, indicating that the extent of the reaction was good but not complete, as expected for a typical Michael reaction addition [(10) and references therein]. MALDIMS analysis performed on the peptide mixture revealed two ions at m/z 1873.5 and m/z 2163.9 assigned to the β-eliminated fragments 106-119 and 104-119, modified by DANS (Δm=310 Da).

In a proof of principle experiment, to investigate the feasibility of applying the method to proteomic analysis, triptic digest from 10 standard proteins (50 μg of a mixture in equimolar amount of myoglobin, bovine serum albumin, ovalbumin, carbonic anhydrase, RNase A, Lysozime, gluthatione S-trasferase, insulin, enolase, trasferrin) was spiked with 5 μg of α-casein triptic mixture modified with dansylcysteamine. 1 pmol of the peptide mixture was then submitted to LCMS analysis by using 4000Q-Trap coupled to a 1100 nano HPLC system. We settled an experiment combining a precursor ion scan with a MS^3 linear ion trap scan modes. The previous one is a typical quadrupole scan mode which can be used when an ion loses a diagnostic fragment as a charged fragment. In a precursor ion scan, Q1 is scanned across a mass range and ions are fragmented in the collision cell. Q3 is set to transmit only the mass of diagnostic fragment. Therefore only ions that are passed through Q1 that fragment to produce the diagnostic charged fragment will be detected. Then the selected precursor ions are submitted to a sequence consisting of MS^2 and MS^3 experiments thus leading to the selective detection of the dansylpeptides. Figure 2A shows the reconstructed ion chromatogram for the selective dansyl transition 234→170 in MS^3 mode. As indicated in the figure, only two ions were detected. Corresponding MS^2 spectra led to the reconstruction of the entire sequences of the beta eliminated phosphopeptides 104-119 and 106-119 carrying a DANSH moiety. As an example, the MS/MS spectrum of the modified peptide 104-119 is reported in Figure 2B. The modified ion is stable during collision induced dissociation, provided easily interpretable product ion spectra. In fact, the y and b product ions still retains the modifying group linked to the beta eliminated Ser residue thus allowing the exact localization of the phosphorylation site. It is worth considering that α-casein contains also a diphosphorylated and a pentaphosphorylated tryptic peptides which escaped to be detected using this approach propably due to the chemical physical characteristics of the modified fragments. In fact the introduction of a high number of dansyl moieties increase the hydrofobicity of the peptides, moreover the low efficiency of Michael addition results dramatically increased when more then one sites are present in the same peptide. However the ionization of multiphosphorylated peptides is a difficult challenge and among the proposed derivatization procedures only the modification strategy based on the use of dithiothreitol (10) led to the detection of intense ions of multiphosphorylated peptides in mass spectrometric analyses.

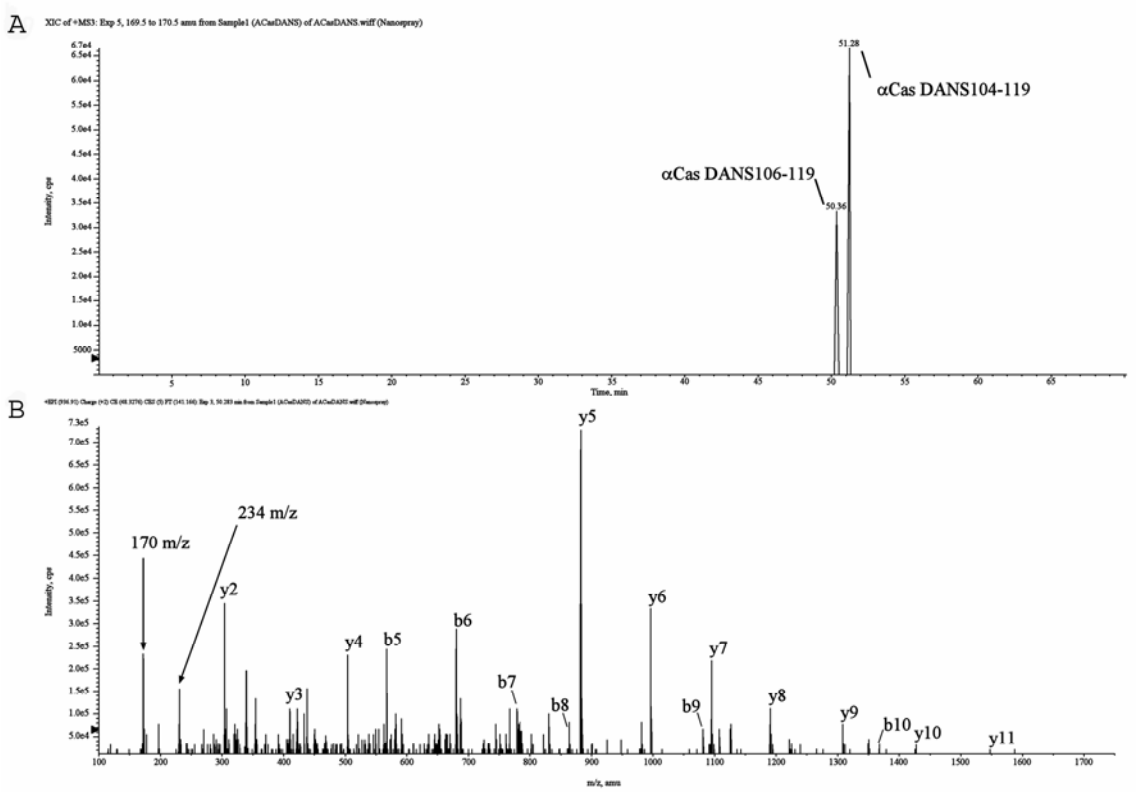


Figure 2. A Reconstructed ion chromatogram of the tryptic digest for the selective transition 234→170 in MS3 mode. B MS/MS spectrum of the α-casein modified peptide 104-119. The fragment ions belonging to the b and y series are indicated

IV.5 Conclusions

In conclusion, it is well known that ion trap tandem mass spectrometry experiments offer high sensitivity because of the ability to accumulate precursor ions. Furthermore, the novel introduction of a linear ion trap with greatly increased capture efficiency and storage capacity resulted in new inputs in the proteomic field. As concerning the use of dansyl chloride as marker, a) DANS moiety introduces a strong basic group in peptides thus leading to an enhanced ionization; b) DANS peptide can be easily detected in precursor ion scan mode; c) DANS peptide can be selectively detected in MS3 mode (via transition 234→170) in even complex peptide mixture. Here the strategy proposed, taking in account the availability of linear trap to select specific labelled peptides giving rise to diagnostic MS^3 product ions, resulted to be of useful in the phosphoprotein analysis. As a whole, this approach will be of interest in the phosphoproteome (and O-glycome) studies and has been successfully exploited to detect selectively phospho-Ser/Thr containing peptides.

IV.6 References

1 Malorni A and Marino G. Peptide fragmentation mechanisms. Adv Mass Spec Biochem Med. 1976. 2:249-255.
2 Addeo F, Malorni A, Marino G., "Analysis of dansyl-1-amides in mixtures by mass spectrometry, using metastable defocusing." Anal Biochem. 1975. 64(1):98-101.

3 Marino G. and Malorni A. Advances in Mass Spectrometry in Biochemistry and Medicine. Spectrum Publications Inc. 1976. vol II. 249.
4 Yan JX, Packer NH, Gooley AA, Williams KLJ. Chromatogr. A. 1998. 808: 23.
5 Schroeder MJ, Shabanowitz J, Schwartz JC, Hunt DF, Coon J.J. Anal. Chem. 2004 76: 3590.
6 Olsen JV, Mann M. Proc Natl Acad Sci U S A. 2004. 101: 13417.
7 Venable JD, Dong MQ, Wohlschlegel J, Dillin A, Yates JR III. Nature Methods. 2004. 1: 1.
8 Marino G and Buonocore V. Biochem J. 1968. 110: 603.
9 Park SJ, Song JS and Kim HJ. Rapid Commun. Mass Spectrom. 2005. 19: 3089.
10 Amoresano A, Marino, G, Cirulli C, Quemeneur E. Eur J Mass Spectrom. 2004. 10: 401.
11 Collings BA, Campbell JM. A combined linear ion trap time-of-flight system with improved performance and MS^n capabilities. Rapid Comm. Mass Spec. 2001.15: 1777–1795.

PUBLICATIONS

1. Selective detection and identification of phosphopeptides by dansyl MS/MS/MS fragmentation.
Amoresano A, Monti G, **Cirulli C**, Marino G.
Rapid Commun Mass Spectrom. 2006. 20:1400-1404.

2. Phosphorylation by protein kinase CK2 modulates the activity of the ATP binding cassette A1 transporter.
Roosbeek S, Peelman F, Verhee A, Labeur C, Caster H, Lensink MF, **Cirulli C**, Grooten J, Cochet C, Vandekerckhove J, Amoresano A, Chimini G, Tavernier J, Rosseneu M.
J Biol Chem. 2004. 279:37779-88.

3. Mapping phosphorylation sites: a new strategy based on the use of isotopically labelled DTT and mass spectrometry. Amoresano A, Marino G, **Cirulli C**, Quemeneur E.
Eur J Mass Spectrom. 2004. 10:401-12.

4. Method to express and purify nm23-H2 protein from baculovirus-infected cells.
Garzia L, Andre A, Amoresano A, D'Angelo A, Martusciello R, **Cirulli C**, Tsurumi T, Marino G, Zollo M.
Biotechniques. 2003. 35:384-8, 390-1.

SCIENTIFIC PRESENTATIONS

1. Phosphoproteomics: a challenge of the post genomic era.
Claudia Cirulli, Angela Amoresano, Matilde Marchetta, Tonia Gamberi, Francesca Magherini, Alessandra Modesti, Guido Camici, Gennaro Marino.
25th Anniversary of Centro Interdipartimentale Di Metologie Chimico-Fisiche Università Di Napoli “Federico II”, Napoli 2003

2. Phosphoproteomics: a challenge of the post genomic era.
Claudia Cirulli, Angela Amoresano, Matilde Marchetta, Tonia Gamberi, Francesca Magherini, Alessandra Modesti , Guido Camici , Gennaro Marino.
Ist National Ihupo Congress, Napoli 2003

3. A Proteomic Approach In The Investigation Of Transcriptional Regulatory Network In E.Coli.
Claudia Cirulli , Angela Amoresano, Nunzianna Doti, and Angela Duilio.
IIst National Ihupo Congress, Chieti 2004.

4. Recognition of pathogen associated molecular patterns (pamps) of gram negative bacteria by mass spectrometry.
Claudia Cirulli; Serena Leone ; Antonio Molinaro; Angela Amoresano.
ISA 2006 , Bari 2006.

5. A proteomic approach in the investigation of transcriptional regulatory network in E.coli.
Claudia Cirulli, Angela Amoresano, Valentina Rippa, Carla Esposito, Nunzianna Doti, Rosanna Papa and Angela Duilio.
Proteine 2006, Novara.

6. Recognition of pathogen associated molecular patterns (PAMPs) of Gram negative bacteria by mass spectrometry.
Claudia Cirulli; Serena Leone; Antonio Molinaro; Angela Amoresano; Piero Pucci.
Itpa 2006, Pisa.

7. Dansyl chemistry in Proteomics: back to the future.
Angela Amoresano, Claudia Cirulli, Giovanni Chiappetta, Gennaro Marino.
Itpa 2006, Pisa.

8. The apoptosis regulator bag3 interacts with the cct chaperone complex.
Birolo L., Tosco A., Cirulli C., Leone A., Pucci P.
Itpa 2006, Pisa.

Appendix

RAPID COMMUNICATIONS IN MASS SPECTROMETRY
Rapid Commun. Mass Spectrom. 2006; **20**: 1400–1404
Published online in Wiley InterScience (www.interscience.wiley.com) DOI: 10.1002/rcm.2461

Selective detection and identification of phosphopeptides by dansyl MS/MS/MS fragmentation

Angela Amoresano*, Gianluca Monti, Claudia Cirulli and Gennaro Marino
Department of Organic Chemistry and Biochemistry, Federico II University of Naples, Naples, Italy

Received 17 November 2005; Revised 22 February 2006; Accepted 27 February 2006

Protein phosphorylation regulates many cellular processes and pathways, such as cell cycle progression, signal transduction cascades and gene expression. Selective detection of phosphopeptides from proteolytic digests is a challenging and highly relevant task in many proteomics applications. Often phosphopeptides are present in small amounts and need selective isolation or enrichment before identification. Here we report a novel approach to label selectively phospho-Ser/-Thr residues by exploiting the features of a novel linear ion trap mass spectrometer. Using dansyl labelling and MS^3 fragmentation, we developed a method useful for the large-scale proteomic profiling of phosphorylation sites. The new residues in the sequence were stable and easily identifiable under general conditions for tandem mass spectrometric sequencing. Copyright © 2006 John Wiley & Sons, Ltd.

Reversible posttranslational modification is a general mechanism for fine-tuning protein structure and function. In particular, protein phosphorylation plays a key role in the regulation of virtually all cellular events. Reversible protein phosphorylation tightly regulates many critical biological processes such as cell cycle, growth, differentiation and metabolism.[1] To gain further insight into the regulation of these processes by reversible phosphorylation, it is often necessary to characterize the phosphorylation state of specific proteins under specific conditions.

Selective detection of phosphopeptides from proteolytic digests is a challenging and highly relevant task in many proteomics applications. Because of the impact of enzyme-catalyzed phosphorylation and dephosphorylation on the function of many proteins,[2] there is a demand to improve the methods for the reliable analysis of phosphopeptides. Phosphoproteins often exist in low abundance in cells alongside the corresponding non-phosphorylated parent forms, and they need to be selectively isolated or enriched before identification. Moreover, phosphopeptides are easily suppressed in mass spectrometry (MS) in both matrix-assisted laser desorption (MALDI) and electrospray (ESI) ionization modes. Furthermore, the low incidence of phosphorylation in proteins makes it important to increase the selectivity of the analysis.[3]

Several groups have reported the analytical potential of chemical derivatization in conjunction with ESI-MS for the analysis of compounds of synthetic and biological origin.[4,5] Treatment with an appropriate derivatizing agent can improve the ionization of a compound and therefore its detectability by ESI-MS. Here we suggest that derivatization may include the addition of a chemical moiety that gives rise to characteristic product ions, thus allowing the exploitation of tandem multi-stage mass spectrometric (MS^n) analysis.[6,7] We introduce the novel acronym RIGhT (reporter ion generating tag) to indicate a general derivatization method based on the labelling of target residues with reagents capable of generating reporter ions in $MS^2 \rightarrow MS^3$ experiments. In the case of phosphopeptides the procedure consists of the β-elimination of phospho-Ser/-Thr residues and the Michael addition of the resulting α,β-unsaturated residues with dansyl-cysteamine. The dansyl derivatization introduces: (a) a basic secondary nitrogen into the molecule that enhances the efficiency of ionization and (b) a dansyl moiety whose fragmentation is known from previous data.[8,9] Using a new hybrid mass spectrometer equipped with a linear ion trap analyzer, one can take advantage of the distinctive m/z 170 and 234 product ions in MS^2 and the diagnostic m/z $234 \rightarrow 170$ fragmentation in the MS^3 mode.

As a whole, this approach, first applied for the identification of phosphorylation sites in model phosphopeptides, promises to enrich our view of the dynamic changes occurring even in complex matrices such as human serum. It is worth considering that O-glycosylated residues can also be targeted with the dansyl reagent thus widening the effectiveness of the method.

**Correspondence to*: A. Amoresano, Dipartimento di Chimica Organica e Biochimica, Complesso Universitario Monte Sant' Angelo, Via Cinthia 6, I-80126 Napoli, Italy.
E-mail: angamor@unina.it
Contract/grant sponsor: Ministero dell'Università e della Ricerca Scientifica (Progetti di Rilevante Interesse Nazionale 2002 and 2003; FIRB 2001); Programma Nazionale di Ricerche in Antartide 2004, Regione Campania L.R. 05/03.

EXPERIMENTAL

Materials and methods

Proteomic grade trypsin, α-casein, myoglobin, bovine serum albumin, insulin, transferrin, ribonuclease A, carbonic anhydrase, lysozyme, glutathione S-transferase, enolase and tributylphosphine solution (stock solution 200 mM in glass vial, sealed under Ar) were purchased from Sigma (St. Louis, MO, USA). Dansyl chloride and 2,2′-dithiobis(ethylamine) (cystamine) were of the highest purity available from Sigma (BioChemika >98.0%). Trifluoroacetic acid (TFA, HPLC grade) was from Carlo Erba (Milano, Italy). All used solvents were of the highest purity available from Baker (Phillipsburg, NJ, USA). All other reagents and proteins were of the highest purity available from Sigma.

Peptides β-elimination

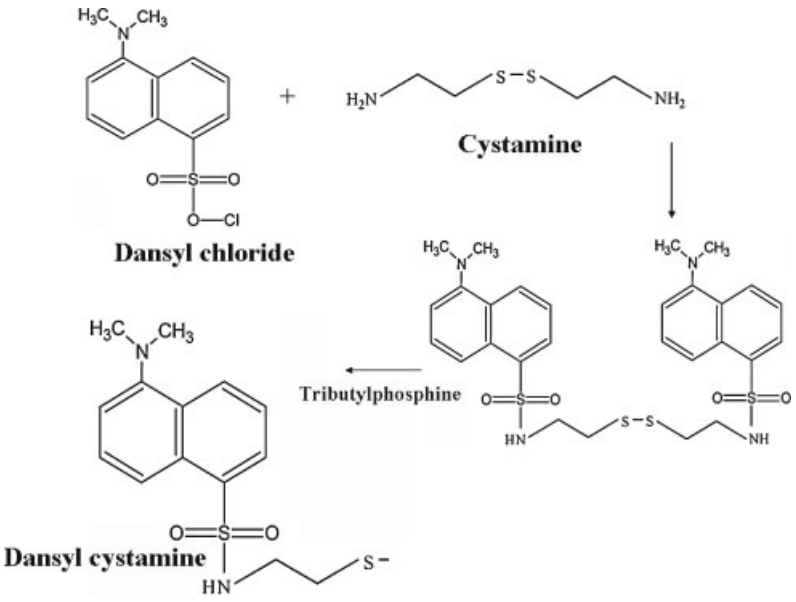
Enzymatic digestion of α-casein (4 nmol/mL) as well as of other standard proteins was carried out using an enzyme/substrate ratio 1:50 (w/w) in a 10 mM ammonium bicarbonate solution, at 37°C for 18 h. The phosphate moieties from a tryptic α-casein digest were removed from pSer and pThr by barium hydroxide ion-mediated β-elimination.The β-elimination reactions were carried out by incubating the peptide mixture in 100 mM $Ba(OH)_2$, at 37°C for 90 min under nitrogen. The reagent excess was removed by using solid carbon dioxide. The precipitated barium carbonate was removed by centrifugation at 13000 rpm for 5 min. The procedure was repeated once. The use of this reagent permitted removal of the barium as a carbonate pellet at the bottom of a vial.[11] The peptide mixture was then analyzed by MALDI-MS.

Dansyl-cysteamine addition

The β-eliminated peptide mixture was modified using dansyl chloride via a Michael-type addition. The dansyl-cysteamine was obtained by reaction of dansyl chloride (0.1 mg/mL dissolved in acetonitrile) with cystamine (molar ratio 3:1) at 4°C for 2 h. The product was purified by reversed-phase high-performance liquid chromatography (RP-HPLC) using an Agilent Zorbax C8 column (4.6 mm × 150 mm; Palo Alto, CA, USA) using a 5% to 65% linear gradient in 30 min from water to acetonitrile (ACN) and verified by ESI-MSMS. Dried dansyl-cysteamine was dissolved in 10 mM Tris/HCl buffer (~15 nmol/μL, pH 8.5) and reduced in the presence of 20 mM tributylphosphine for 15 min at room temperature under nitrogen (see Scheme 1). The product was purified by RP-HPLC as previously described and verified by ESI-MS/MS. Then, 100 μL of dansyl-cysteamine solution (0.1 mg/mL, dissolved in ACN/Water 3:1) were added to an equal volume of the β-eliminated peptide mixture solution (1 nmol/mL, in deionized water) in the presence of 1 mM tributylphosphine. Tris (2 M) was added to the sample at a final concentration of 100 mM (pH ~11). The reaction proceeded under nitrogen for 18 h at 50°C.

MALDI-TOFMS

MALDI-TOF mass spectra were recorded using a Voyager DE-PRO mass spectrometer (Applied Biosystems, Framingham, MA, USA). The peptide mixtures were first purified using ZipTip pipettes from Millipore (Billerica, MA, USA) using the manufacturer's recommended purification procedure. The peptides were eluted using 20 μL of 50% ACN, 0.1% TFA in water. A mixture of the eluted peptide solution and α-cyanohydroxycinnamic acid (10 mg/mL in ACN/0.1% TFA in water, 2:1, v/v) was applied to the metal sample plate and dried at room temperature. Mass calibration was performed using a mixture of peptides from Applied Biosystems, containing des-Arg1-bradykinin, angiotensin I, Glu1-fibrinopeptide B, ACTH (1–17), ACTH (18–39) and insulin (bovine). Raw data were analyzed using Data Explorer software provided by Applied Biosystems and reported as monoisotopic masses.



Scheme 1. Dansyl-cysteamine synthesis.

DOI: 10.1002/rcm

Nano-LC/MS

The peptide mixture was analyzed by LC/MS using a 4000Q-Trap (Applied Biosystems) equipped with a linear ion trap coupled to an 1100 nano-HPLC system (Agilent Technologies). The mixture was loaded on an Agilent reversed-phase pre-column cartridge (Zorbax 300 SB-C18, 5 × 0.3 mm, 5 μm) at 10 μL/min (solvent A: 0.1% formic acid, loading time 7 min). The peptides were separated on a Agilent reversed-phase column (Zorbax 300 SB-C18, 150 mm × 75 μm, 3.5 μm), at a flow rate of 0.2 μL/min with a 5 to 65% linear gradient in 60 min (solvent A: 0.1% formic acid, 2% ACN in water; solvent B: 0.1% formic acid, 2% water in ACN). A micro-ionspray source was used at 2.4 kV with liquid coupling, with a declustering potential of 50 V, using an uncoated silica tip (o.d. 150 μm, i.d. 20 μm, tip diameter 10 μm) from New Objectives (Ringoes, NJ, USA). Spectra acquisition was based on a survey precursor ion scan for *m/z* 170. The Q1 quadrupole was scanned from *m/z* 500–1000 in 2 s, and the precursor ions were fragmented in q2 using a linear gradient of collision potential from 30 to 70 V. Finally, Q3 was set to transmit only ions at *m/z* 170. This scan mode was followed by an enhanced resolution experiment (ER) for the ions of interest and then by MS^3 and MS^2 acquisitions of the two most abundant ions. MS^2 spectra were acquired using the best collision energy calculated on the bases of *m/z* values and charge state (rolling collision energy). MS^3 spectra were acquired using Q0 trapping, with a trapping time of 150 ms and an activation time of 100 ms, scanning from *m/z* 160–240. The Q0 trapping facility is able to increase the sensitivity of the scan in the different ion trap scan types used. Data were acquired and processed using Analyst software (Applied Biosystems).

RESULTS AND DISCUSSION

In an attempt to improve and simplify the *plethora* of methods so far suggested[11] we have set up a new approach to label selectively phospho-Ser/-Thr residues by exploiting the features of a novel linear ion trap mass spectrometer.[12] Dansyl chloride is known to react with the ε-amino group of lysine as well as with the N-terminal amino group.[10] Here, using chemical manipulations coupled with dansyl chloride labelling, we show a methodology capable of large-scale proteomic profiling of phosphorylation sites. The method is based on (a) the selective modification of pSer/Thr residues with a novel dansyl derivative of cysteamine, prepared as reported in the Experimental section and shown in Scheme 1, and (b) the selective detection and identification of labelled peptides by exploiting the characteristic fragmentation pathway of dansyl derivatives.

We first applied this approach to α-casein as a model phosphoprotein. The phosphate moieties from a tryptic α-casein digest were removed from pSer and pThr by barium hydroxide ion-mediated β-elimination as already reported.[11] The peptide mixture was then analyzed by MALDI-MS, using the Voyager DE-PRO mass spectrometer.

The MALDI mass spectra showed ions corresponding to the β-eliminated peptides. As an example, the ion at *m/z* 1854.3 was assigned to the peptide 104–119 of α-casein occurring 98 Th lower than expected, indicating that the pSer115 was converted into dehydroalanyl after the β-elimination reaction (Fig. 1). The α,β-unsaturated residues are Michael acceptors, which can readily react with a nucleophile. The β-eliminated peptide mixture was modified with dansyl-cysteamine via Michael-type addition.

The yield of the addition reaction was monitored by MALDI-MS, and indicated that the extent of the reaction was good but not complete, as expected for a typical Michael reaction addition (see Ref. 11, and references cited therein). MALDI-MS analysis performed on the peptide mixture revealed two ions at *m/z* 1873.5 and 2163.9 assigned to the β-eliminated fragments 106–119 and 104–119, modified by a dansyl moiety ($\Delta m = 310$ Da).

In a proof-of-principle experiment, to investigate the feasibility of applying the method to proteomics analysis, a tryptic digest from ten standard proteins (50 μg of a mixture in equimolar amounts of myoglobin, bovine serum albumin, ovalbumin, carbonic anhydrase, RNase A, lysozyme, glutathione S-transferase, insulin, enolase, transferrin) was spiked with 5 μg of an α-casein tryptic mixture modified with dansyl-cysteamine. Then 1 pmol of the peptide mixture was analyzed by LC/MS using the 4000Q-Trap coupled to the 1100 nano-HPLC system. We set up an experiment combining a precursor ion scan with an MS^3 linear ion trap scan. In the precursor ion scan mode, Q1 is scanned across the full mass range, and ions are fragmented in the collision cell. Q3 is set to transmit only the mass of the diagnostic product ion (*m/z* 170). Therefore, only those precursor ions that pass through Q1 that fragment to produce *m/z* 170 are detected. The selected precursor ions are then submitted to a combined MS^2 and MS^3 experiment to specifically detect only those ions that produce the transition *m/z* 234 → 170 in the MS^3 scan mode leading to the selective detection of the dansylpeptides. Figure 2(A) shows a typical reconstructed ion chromatogram for the transition *m/z* 234 → 170 in MS^3 mode. As indicated in the figure, only two ions were detected. Corresponding MS^2 spectra led to the reconstruction of the entire sequences of the β-eliminated phosphopeptides 104–119 and 106–119 carrying a dansyl-cysteamine moiety. As an example, the MS/MS spectrum of the modified peptide 104–119 is reported in Fig. 2(B). The modified ion is stable during collision-induced dissociation, providing easily interpretable product ion spectra. In fact, the y and b product ions still retain the modifying group linked to the β-eliminated Ser residue, thus allowing the exact localization of the phosphorylation site. However, this approach seems to be amenable only for monophosphorylated peptides since we were unable to detect the diphosphorylated and pentaphosphorylated α-casein tryptic peptides. The detection of multiphosphorylated peptides is a difficult challenge and, among the proposed derivatization procedures, to the best of our knowledge, only the modification strategy that we have proposed based on the use of dithiothreitol[11] has resulted in the confident mass spectrometric analysis of multiphosphorylated peptides.

DOI: 10.1002/rcm

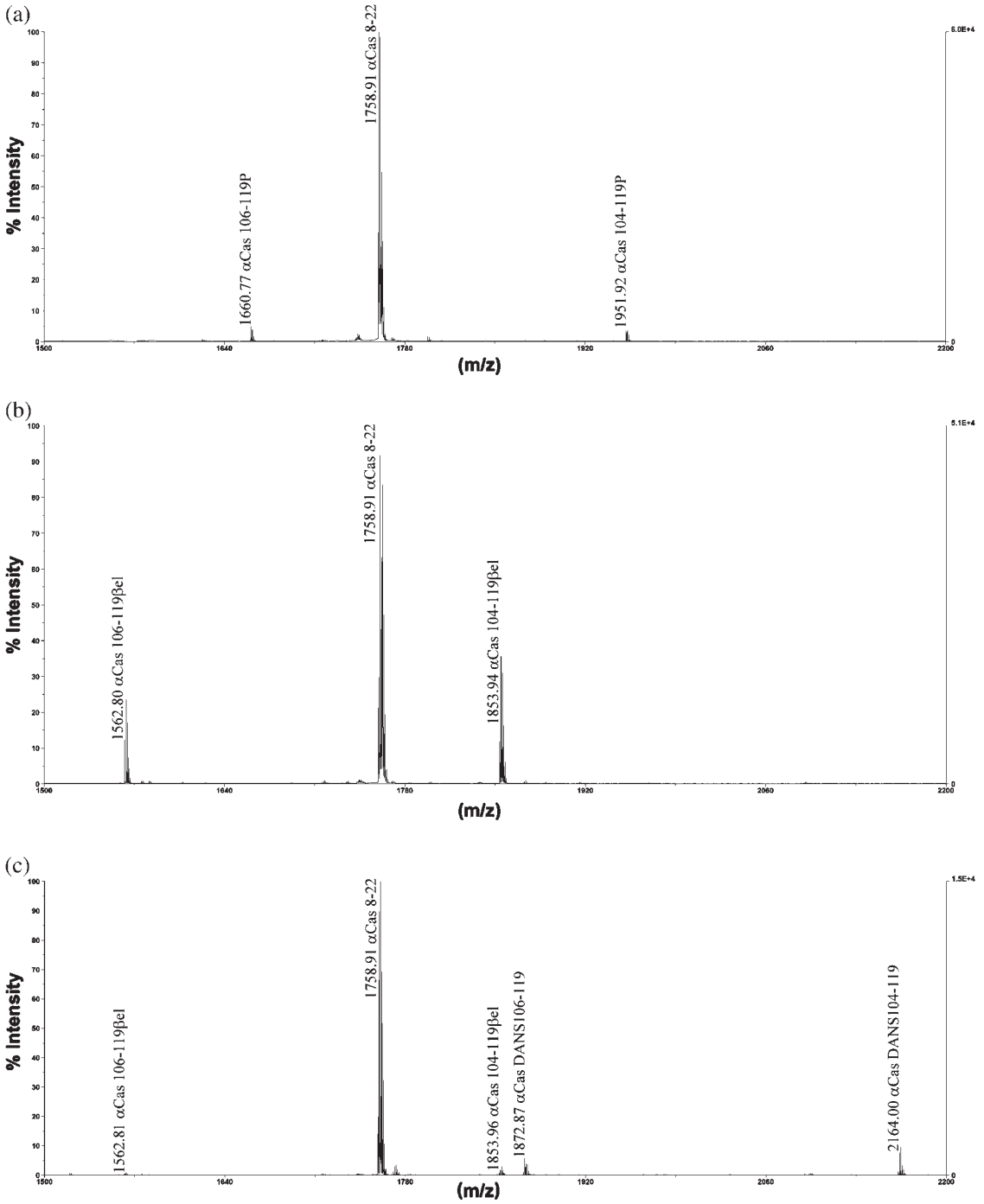


Figure 1. (a) Partial MALDI mass spectrum of the trypsin digest of α-casein. (b) Partial MALDI mass spectrum of the β-eliminated peptide mixture. The ions related to the phosphorylated peptides are shifted by 98 Th due to the β-elimination reaction. (c) Partial MALDI mass spectrum of the β-eliminated peptide mixture after dansyl-cysteamine addition. The ions corresponding to the β-eliminated peptides show a mass increment of 310 Da.

CONCLUSIONS

It is well known that ion trap tandem mass spectrometry offers high sensitivity because of the ability to accumulate precursor ions. Furthermore, the introduction of a linear ion trap with greatly increased capture efficiency and storage capacity has resulted in new inputs in the proteomics field. Use of dansyl chloride as marker leads to (a) the dansyl moiety introducing a strong basic group in peptides thus giving enhanced ionization; (b) easy detection of dansyl peptides in the precursor ion scan mode; and (c) selective detection of dansyl peptides in the MS^3 mode (via the transition m/z 234 $\rightarrow$ 170) even in a complex peptide mixture. The proposed strategy, using a linear ion trap to select specific labelled peptides giving rise to diagnostic MS^3 product ions, was useful in phosphoprotein analysis. This

DOI: 10.1002/rcm

RCM

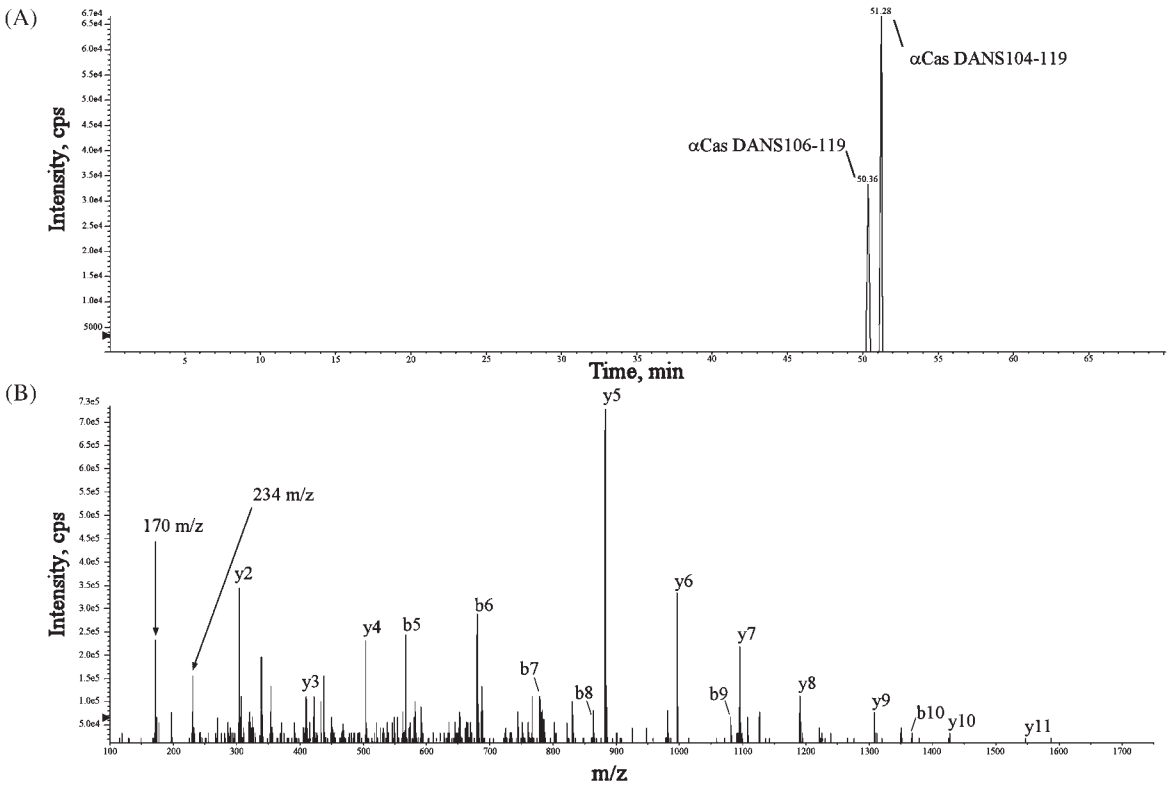


Figure 2. (A) Reconstructed ion chromatogram of the tryptic digest for the selective transition *m/z* 234 → 170 in MS^3 mode. (B) MS/MS spectrum of the α-casein modified peptide 104–119. The product ions belonging to the b and y series are indicated.

approach will be of interest in phosphoproteome (and O-glycome) studies and it has been successfully exploited to selectively detect phospho-Ser/-Thr-containing peptides. The introduction of MS^3 analysis actually replaces an off-line chromatographic step (affinity chromatography) aimed at isolating the labelled peptides (see Ref. 11, and references cited therein) with the obvious advantages of avoiding time-consuming analysis and increasing sensitivity.

Acknowledgements

This work was supported by grants from Ministero dell'Università e della Ricerca Scientifica (Progetti di Rilevante Interesse Nazionale 2002 and 2003; FIRB 2001), Programma Nazionale di Ricerche in Antartide 2004, and Regione Campania L.R. 05/03. Support from the National Center of Excellence in Molecular Medicine (MIUR—Rome) and from the Regional Center of Competence (CRdC ATIBB, Regione Campania—Naples) is gratefully acknowledged.

REFERENCES

1. Hunter T. *Cell* 2000; **100**: 113.
2. Yan JX, Packer NH, Gooley AA, Williams KL. *J. Chromatogr. A* 1998; **808**: 23.
3. Schroeder MJ, Shabanowitz J, Schwartz JC, Hunt DF, Coon JJ. *Anal. Chem*. 2004; **76**: 3590.
4. Dalvie DK, O'Donnell JP. *Rapid Commun. Mass Spectrom*. 1998; **12**: 419.
5. Pearson PG, Threadgill MD, Howald WN, Baillie TA. *Biomed. Environ. Mass Spectrom*. 1988; **16**: 51.
6. Olsen JV, Mann M. *Proc. Natl. Acad. Sci. USA* 2004; **101**: 13417.
7. Venable JD, Dong MQ, Wohlschlegel J, Dillin A, Yates JR III. *Nat. Method* 2004; **1**: 1.
8. Marino G, Malorni A. Advances in mass spectrometry. In *Biochemistry and Medicine*, vol II. Spectrum Publications. Inc.: New York, 1976; 249.
9. Marino G, Buonocore V. *Biochem. J*. 1968; **110**: 603.
10. Park SJ, Song JS, Kim HJ. *Rapid Commun. Mass Spectrom*. 2005; **19**: 3089.
11. Amoresano A, Marino G, Cirulli C, Quemeneur E. *Eur. J. Mass Spectrom*. 2004; **10**: 401.
12. Collings BA, Campbell JM. *Rapid Commun. Mass Specrom*. 2001; **15**: 1777.

DOI: 10.1002/rcm

Method to express and purify nm23-H2 protein from baculovirus-infected cells

L. Garzia[1], A. André[1], A. Amoresano[2], A. D'Angelo[1], R. Martusciello[1], C. Cirulli[2], T. Tsurumi[3], G. Marino[2], and M. Zollo[1]

[1]*Telethon Institute of Genetics and Medicine, Naples, Italy,* [2]*School of Biotechnological Sciences, Università di Napoli Federico II, Naples, Italy, and* [3]*Aichi Cancer Center Research Institute Division of Virology, Nagoya, Japan*

BioTechniques 35:384-391 (August 2003)

High-throughput protein expression and purification are major bottlenecks in the postgenomic and proteomic era. We show here an automated method to express and purify nm23-H2, a nucleoside diphosphate kinase (NDPK), in a 96-well format, by the use of a robotic workstation, from insect Spodoptera frugiperda *(Sf9) baculovirus-infected cells using nickel-nitrilotriacetic acid (Ni-NTA) agarose beads. The automated method is coupled to mass spectrometry for a validation and quality-control analysis. To verify the bona fide of the recombinant protein, several tests have been produced, including NDPK assay, Western blotting, and in vitro phosphorylation experiments, thus confirming the value of the protocol developed. The method has been validated for the expression of several proteins, thus confirming the value of this automated protocol. The research presented here is a useful method both for industrial and academic environments to produce in a high-throughput mode recombinant eukaryotic proteins to be assayed for a specific function in a systematic manner.*

INTRODUCTION

With the evolution of genome projects and the improvements in DNA sequencing technology, several organisms sequenced have been deposited in public domain databases [i.e., Archaea, bacteria, eukaryotes, viruses, and organelles; see the National Center for Biotechnology Information (NCBI) Web site at http://www.ncbi.nlm.nih.gov/]. The vast number of proteins predicted and/or verified by comparing with a cDNA database is creating massive opportunities for both basic and applied research. One of the major efforts in the postgenomic and functional genomic era is to understand how unknown gene products cooperate in cellular environments to exert their cellular function. Efficient and rapid expression of genes in homologous/heterologous protein expression systems and rapid purification steps are today the major bottlenecks.

To date, methods have been generated to express simultaneously a high number of proteins, based on *Escherichia coli* and *Saccharomyces cerevisiae* expression systems. Both of these systems are adapted to grow and express proteins, in a 96-well format, to obtain partially or fully automated processes. The baculovirus expression system has not yet been tested for its usefulness to express protein in this format, mainly because of its low efficiency on a 96-well-format growth and its incapacity to obtain a reasonable amount of purified protein. Nevertheless, a high yield in a single preparative heterologous expression system has been reported to date. As far as the novel eukaryotic gene expression systems are concerned, there are many advantages in using baculovirus for heterologous gene expression (1). As with other eukaryotic expression systems, baculovirus expression of heterologous genes permits folding, posttranslational modification, and oligomerization in manners that are close, often identical, to those that occur in mammalian cells. Low levels of expression can often be increased with the optimization of time of expression and multiplicity of infection (MOI). Very recently, the use of molecular chaperones (2), ready cloned in baculovirus vectors, has completely encompassed the potential misfolding and solubility problems in insect cells (3).

To investigate if this system should be used in experiments that include the semi-automated expression of a high number of proteins, we evaluated the yield and the quality of the protein expression and purification procedure carried out in a 96-well format.

The recombinant baculovirus used in this work encoded for nm23-H2, a protein for which a nucleoside diphosphate kinase (NDPK) and hystidine phosphotransferase characterized activity, a potential casein kinase II phosphorylation site (4), and a DNA binding and a c-*Myc* transactivating activity were found in vitro and in vivo (5–7).

Today there is an increasing demand for automated protein purification because, to date, the above step is widely considered to be a bottleneck in the assembly line of recombinant protein production. Automation would actually allow one to avoid repeating column chromatographic steps and, as an added benefit, also allows the standardization of the amount of protein produced. This issue has been focused on the *E. coli* expression environment, suggesting the use of nickel-nitrilotriacetic acid (Ni-NTA)-coated multi-well microplates (8). The present state of the art in our laboratory is based on the use of Ni-NTA magnetic agarose beads and their affinity to 6×His-tag protein directly isolated from the lysis buffer environment as presented by Lanio et al. (8). Compared to this method, the advantage of using agarose magnetic beads versus coated plates is the overall increased protein maximum binding capacity of the beads (0.2 nmol agarose Ni-NTA beads/well vs. 10 pmol coated Ni-NTA plates/well), thus resulting in a significant reduction in the purification costs. A second advantage is that the method allows host growing and several purification steps in the same multi-well plate. Finally, the magnetic agarose NI-NTA beads can be used at least twice without affecting the protein purification rate.

Automation is not meant to obtain large quantities of protein but rather to produce automated simultaneous purification processes for several samples, such as purifying proteins in a 96-well plate simultaneously (a useful method for protein mutational analysis). The following are further advantages of using such an approach: (*i*) the sonication steps, often required to cleave genomic DNA, can be avoided; (*ii*) the fishing of

6× His proteins from cell lysates can be iterated many times to saturate the beads, thus improving their loading capacity; (*iii*) the method has been implemented with magnetic robotic stations. Indeed, the latter devices have already been introduced in the case of DNA purification in the assessment of the human genome draft sequence and can eventually be implemented in automating the protein purification, too.

Appended to the purification step, a quality-control analysis of recombinant products has been implemented by the use of direct analysis of protein-loaded beads using matrix-assisted laser desorption ionization (MALDI) mass spectrometry. Indeed, mass spectrometry plays an important role in proteomics, particularly in protein identification analysis, but also in the characterization of posttranslational modifications. In conjunction with classical protein and carbohydrate chemistry, these strategies have proved to be powerful and cost-effective methods for structural analysis, thus playing a key role in the quality control of recombinant proteins and glycoproteins (9–11). The first step in the elucidation of protein structure generally consists in the determination of the accurate molecular weight of the intact protein. The accuracy of the molecular mass determination confirms, or otherwise the encoded protein sequence suggests, the presence of unexpected covalent modifications. The nature and fine location of any structural modification in the protein sequence can then be identified and assigned using the mass-mapping strategy. Errors of translation, deletion, insertion, point mutation (except inversions in the sequence), posttranslational modifications or processing, and the S-S bridge pattern can be detected and assigned using this method. The sites and the nature of modifying groups originated by post-biosynthetic events (disulfide bonds, methylation, phosphorylation, hydroxylation, or glycosylation) can be precisely assessed. Some strategies have been specifically developed to characterize both N- and O-linked glycans, which integrate chemical manipulations, enzymatic degradation, and mass spectrometry analytical techniques. The definition of the entire structure of the oligosaccharide moieties released from glycoproteins is achieved by determining the microheterogeneity of the different structures, the composition of the monosaccharides present, the branching pattern, the sequence of the antennae, and the possible occurrence of modifying groups, such as sulfate, phosphate, acetyl groups, etc. (12).

MATERIALS AND METHODS

Cell Culture and Baculovirus Stock Preparation

Spodoptera frugiperda (Sf9) cells were cultured in TNM-FH media (Sigma, Milan, Italy) with 10% fetal bovine serum (FBS), 100 U/mL penicillin, and 100 µg/mL streptomycin at 27°C. The cDNA for nm23-H2 was cloned in pFastBacHTa (Invitrogen, Carlsbad, CA, USA), and the recombinant baculovirus was obtained according to the manufacturer's instructions. With a cycle of amplification in the Sf9 cell at an MOI of 0.1, we obtained a virus stock preparation with an approximate titer of 2×10^7 pfu/mL.

Automated Protocol for Protein Expression and Purification

Sf9 cells were plated at a density of 5×10^5/mL in a 96-well flat-bottom plate (Qiagen, Hilden, Germany) and infected at an MOI of 2 and 5 for 48 h at 27°C. The recombinant protein was first purified by Ni-NTA agarose magnetic beads (Qiagen) using 20 µL 5% suspension directly added to the crude lysate, manually applying the method created for the liquid handling workstation. Protein expression and purification under this condition were verified by Western blot analysis following a standard protocol. For the execution of the semi-automated method with a liquid handling workstation, Sf9 cells were grown using the condition previously optimized for the culture in the 96-well format. The infection was performed at an MOI of 2. After 48 h, the cells were treated by the liquid handling workstation (MultiPROBE® II HT-EX; Packard Instruments; PerkinElmer Life Sciences, Boston, MA, USA), the supernatant (150 µL) was removed, and the cell was lysed by 100 µL lysis buffer (50 mM NaH_2PO_4, pH 8.0, 300 mM NaCl, 10 mM imidazole, 0.005% Tween® 20) by up-and-down pipetting with conductive disposable tips (Molecular BioProducts, San Diego, CA, USA). The lysates were then supplemented with 20 µL 5% suspension of Ni-NTA agarose magnetic beads and incubated at 10°C on an Eppendorf® Thermomixer Comfort (Eppendorf, Hamburg, Germany) at 500 rpm for 2 h. The 96-well plate was then placed on the type-A magnet (Qiagen), and the flow-through was then aspirated from the lysis plate with conductive disposable tips and pipetted in a replica plate, which was stored for successive analyses. The lysis plate was removed from the magnet, and the beads were resuspended and washed with 150 µL wash buffer (50 mM NaH_2PO_4, pH 8.0, 300 mM NaCl, 20 mM imidazole, 0.005% Tween 20) by up-and-down pipetting using the syringe-mode configuration. This washing step was repeated twice; in the last step, 150 µL of the bead suspension were transferred in a round-bottom 96-well plate to increase protein recovery. Then, the recombinant protein was eluted, incubating the beads in 50 µL elution buffer (50 mM NaH_2PO_4, pH 8.0, 300 mM NaCl, 250 mM imidazole, 0.005% Tween 20) on the shaker for 5 min at 10°C. The round-bottom plate was then positioned on the magnet tool for 5 min, and the eluates were collected in a replica plate for further analysis on the thermomixer block. Accessory material describing detailed robotic settings and scripts ready to run on the Packard robot are available at (http://seqcore.tigem.it/baculo).

Protein Quantification

The recombinant protein obtained was quantified according to the Bradford method directly in the replica plate obtained from the purification step, using a 96-well ELISA microplate reader (model 550; Bio-Rad Laboratories, Hercules, CA, USA) with bovine serum albumin (BSA) as standard.

Western Blot Analysis

Elution fraction (35 µL) obtained from the protein purification protocol was loaded on 12.5% (w/v) polyacrylamide gel for sodium dodecyl sulfate polyacrylamide gel electrophoresis

(SDS-PAGE) and electroblotted onto a polyvinylidene difluoride (PVDF) membrane (Immobilon™-P; Millipore, Bedford, MA, USA). Monoclonal antibody penta-HIS (Qiagen) anti-HIS epitope was used according to the manufacturer's protocol to detect insect-expressed proteins. The proteins were further analyzed by a polyclonal antibody (Novocastra Laboratories, Peterborough, UK) that recognizes nm23-H2; this antibody was used at 1:500 dilution.

NDPK Assay

We used the method described previously (13), using 100 ng chromatography resin-purified proteins nm23-H1 and nm23-H2. The control killing activity mutation in the nm23-H2H118F protein was made as described previously (14). The slope of a linear plot of absorbance versus time was used as a measure of enzymatic activity.

In Vitro Protein Phosphorylation

Intact protein immobilized on magnetic beads was incubated in 1× reaction buffer (50 µL) supplemented with 200 µM ATP for 1 h at 30°C. Two units of recombinant casein kinase I (New England Biolabs, Beverly, MA, USA) were used. Casein kinase I was then eliminated with three washes with double-distilled water; the beads with the intact phosphorylated protein were resuspended in double-distilled water and immediately analyzed by mass spectrometry.

Mass Spectrometry Analysis

MALDI mass spectrometry analyses were carried out using a linear Voyager™ DE or a reflectron Voyager DE-PRO mass spectrometer (Applied Biosystems, Foster City, CA, USA) equipped with the delayed extraction device. Mass calibration was performed with apomyoglobin at 16951.5 Da and BSA at 66431.4 Da as internal standards for intact protein and with bovine insulin (average molecular mass 5734.6 Da) and a matrix peak (379.1 Da) for the tryptic mixtures. Ten microliters of 0.1% trifluoroacetic acid (TFA) were

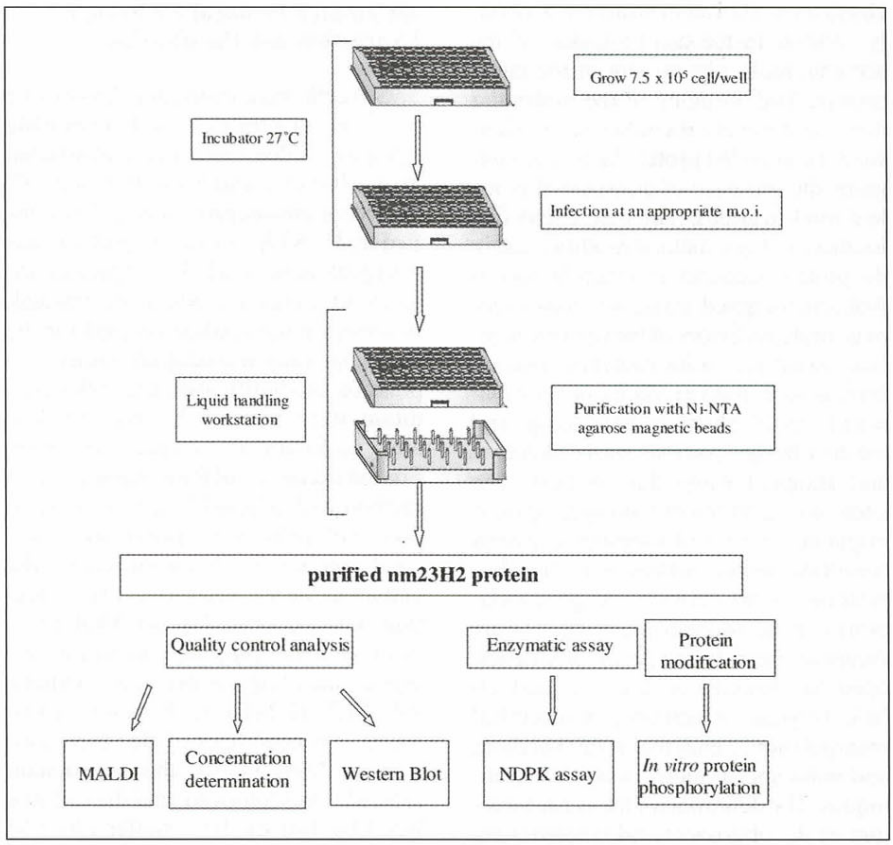


Figure 1. Method describing growth, infection, and purification of nm23-H2 protein.

added to the beads, and 1 µL bead suspension was mixed with 1 µL sinapinic acid (25 mg/mL) in CH_3CN/0.2% TFA 70:30 (v/v) and air-dried. Fifty microliters of 50 mM ammonium bicarbonate, pH 8.5, were directly added to the bead suspension. Trypsin hydrolysis of the protein was carried out at 37°C for 18 h using 1:50 (w/w) enzyme-to-substrate ratio. Tryptic mixtures were dissolved into a final concentration of 10 µM, and 1 µL was applied to a sample slide and mixed with 10 mg/mL α-cyano-4-hydroxycinnamic acid solution in CH_3CN/0.2% TFA 70:30 (v/v) before air-drying. Mass spectra were generated from the sum of 500 laser shots.

RESULTS AND DISCUSSION

We at first defined the best MOI to infect the Sf9. At an MOI of 2, the protein production reached the plateau as confirmed by Western blot analysis (see Figure 2). The same conditions were applied to the use of chaperone proteins described previously (10). Co-infection with baculovirus chaperons Hsp40 and Hsp70 increased protein folding, resulting in 50 ng protein/well (data not shown). We used 20 µL magnetic bead suspension; this amount has a binding capacity of 0.2 nmol, largely exceeding the quantity of the recombinant protein expected from previous preparative experiments. The semi-automated purification procedure was performed from a 96-well plate infected for 48 h at an MOI of 2.

After the identification of the right MOI infection conditions, we have infected 96-well plate containing 5×10^5 seed Sf9 cells in TNM-FH insect cell-supplemented media. After 48 h incubation at 27°C, we have isolated the multi-well and performed the purification on the robotic workstation (Figure 1). The method and conditions totaled 3 h. The purified protein product for each well was measured by a colorimetric assay and evaluated as 50 ng protein purified/well, totaling almost 5 µg puri-

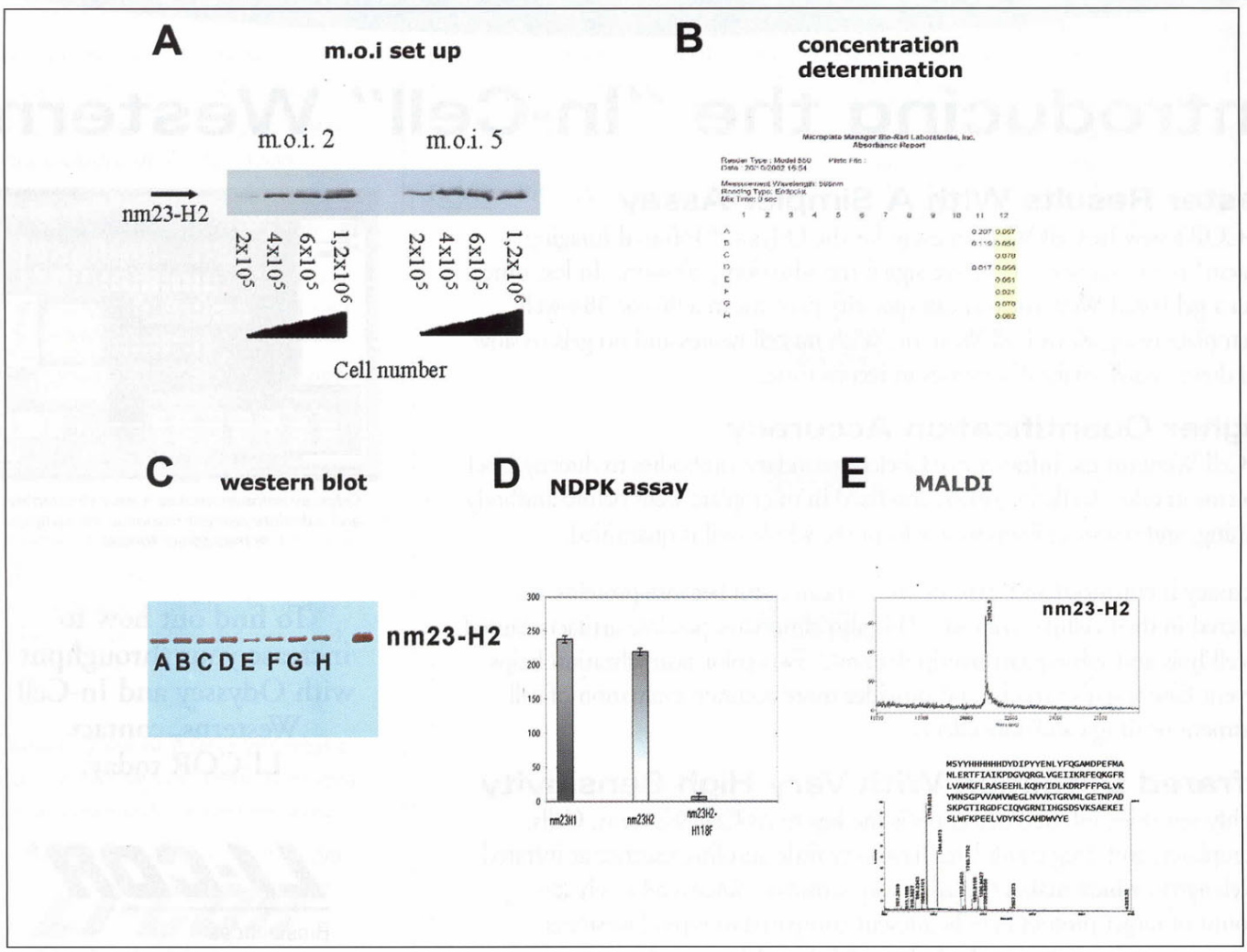


Figure 2. Validation of the protein purification protocol. (A) Identification of the correct multiplicity of infection (MOI) of Sf9 cells, ranging from 2×10^5 to 1.2×10^6 total number of cells, using two different virus concentrations (MOI of 2 and 5). (B) Protein quantity determination performed on randomly selected wells, from the 96-well plate acquired by the method. (C) Western blot analysis using polyclonal antibodies (anti-nm23-H2) used at 1:500 dilution. As positive control, 50 ng nm23-H2 purified protein were used. (D) Nucleoside diphosphate kinase (NDPK) assay to determine the specific activity of recombinant protein obtained from the method. As a negative control, a killer catalytic site mutant (nm23-H118F) was used.The y-axis shows U/mg protein identified. (E) Protein spectrum by MALDI analyses directly obtained from recombinant protein immobilized on magnetic agarose nickel-nitrilotriacetic acid beads.

fied protein/96-well plate.

Following the steps presented in Figure 1, the protein preparation was checked for activity (NDPK assay), identity (Western blot analysis and MALDI mass spectrometry), and post-translational modifications (tryptic digestion and MALDI mass spectrometry). Through Western blot analysis (Figure 2, panels A and C), we did not observe any product derived from an incomplete protein synthesis or any degradation products.

Because the protein produced is a nucleoside diphosphate kinase, we applied the assay to verify its activity. The NDPK assay applied is the same presented in Reference 14. The nm23-

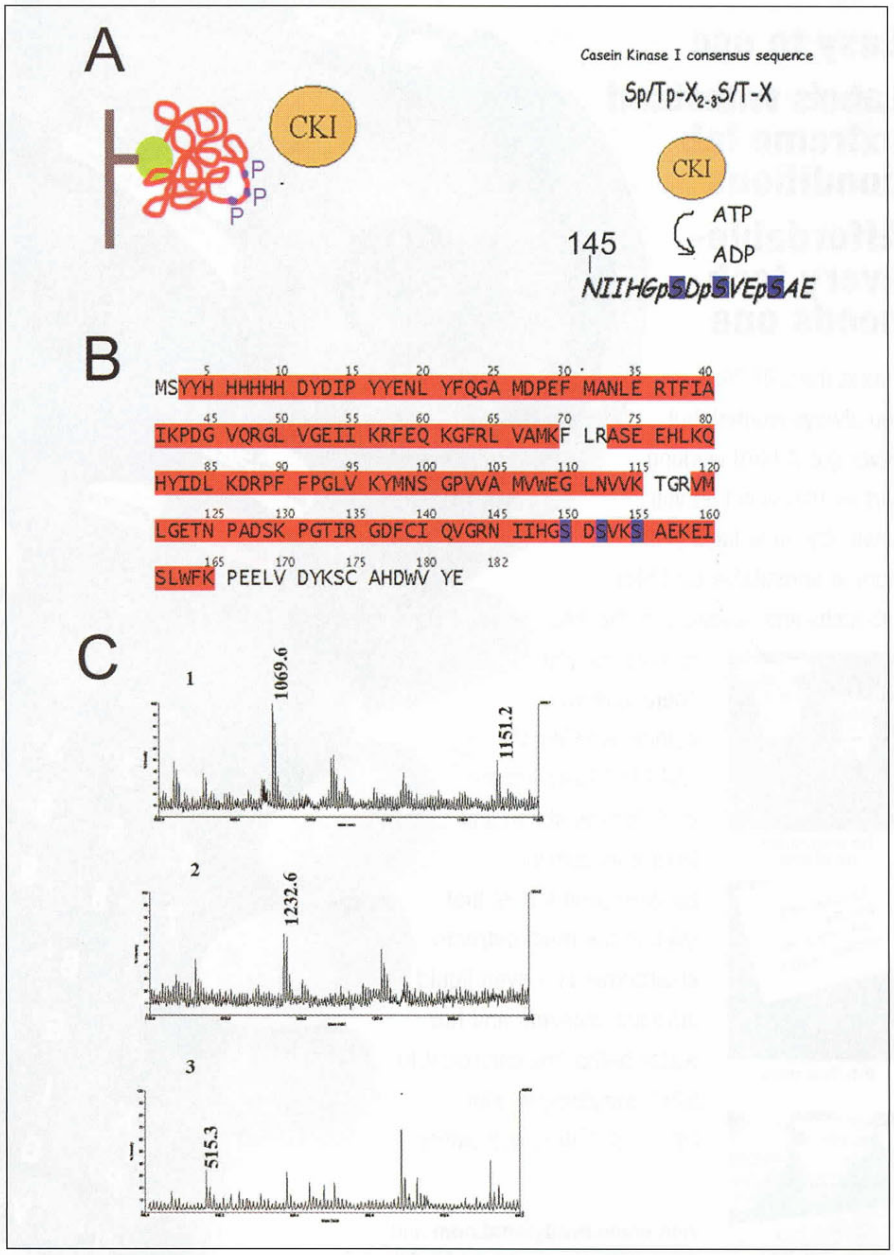


Figure 3. In vitro phosphorylation and MALDI mass spectrometry analyses of nm23-H2 purified protein. (A) A graphic representation of His-nm23-H2 protein attached to agarose nickel-nitrilotriacetic acid beads and a casein kinase I (CKI) protein motif recognition site of phosphorylation on nm23-H2 protein. (B) In red, protein sequence obtained by tryptic digestion resolved by MALDI; white sequences correspond to regions not solved. In blue, serines 120, 122, and 125 found phosphorylated. (C) Partial MALDI mass spectrometry spectrum of the tryptic digest nm23-H2 phosphoprotein. The phosphorylated peptides at m/z 1069.6 and 1151.5 (1), 1232.6 (2), and 515.3 (3) are reported in the panels.

H2 protein produced is active as an NDPK 218 U/mg; as a negative control, we made a second protein nm23-H2-H118F mutation, killing the catalytic activity (Figure 2D).

Intact protein immobilized into magnetic beads was directly analyzed by mass spectrometry to obtain information on the exact molecular mass. The protein was subjected to MALDI mass spectrometry, producing the spectrum shown in Figure 2E. A single sharp peak was detected in the mass spectrum, exhibiting an average molecular mass of 20920.3 Da. An aliquot of nm23-H2 was then digested with trypsin, and the resulting peptide mixture was directly analyzed by MALDI mass spectrometry, producing the spectrum shown in Figure 3C. The mass spectral analysis led to the verification of about 92% of the entire protein sequence (Figure 3B).

An additional assay was performed to determine if nm23-H2 in vitro can be phosphorylated by casein kinase I. A previous article (4) has already shown the potential of nm23 protein to be phosphorylated by casein kinase II.

In addition, using the predicting phosphorylated software available at (http://www.cbs.dtu.dk/services/NetPhos/ and http://www.cbs.dtu.dk/databases/PhosphoBase/), we have observed a region around S120–S125 that could be a potential site for casein kinase I (Figure 3A).

The MALDI analyses showed the presence of three series of related peaks (Figure 3C). The other two series of doublets were assigned to the mono- and diphosphorylated forms of the protein isoforms. Serines 125, 122, and 120 were found phosphorylated as shown in trypsin digestion pattern (Figure 3C).

In this report, we show our capability to express nm23-H2 protein (20 kDa in size) in a 96-well format and purify it by using Ni-NTA agarose beads all the way by using a robotic workstation, developing an automated method coupled to mass spectrometry within a detailed quality-control analysis. To guarantee and increase the rate of purified protein, the method presented here uses chaperones recombinant (Hsp40 and Hsp70) baculovirus co-infection to ensure appropriate protein folding. The automation procedure runs in 3 h and is a useful method for industrial and academic environments to produce large-scale phase-recombinant eukaryotic proteins to be assayed for function in a systematic manner. To date, our laboratory has used this method to purify several expressed proteins ranging from 20 to 70 kDa in size, with a recovery rate of 2–4 pmol purified protein/well, thus further confirming the reproducibility and consistency of the automated protocol.

In addition, this method will allow one to characterize the catalytic domain and/or the specific interaction sites in a protein complex; experiments described as of random and site directed mutagenesis, generating high number of mutants, would be expressed at the same time and tested for their enzymatic activity. Finally, several tests have been performed to verify the quality of the recombinant protein produced, including NDPK activity assays and in vitro phosphorylation experiments to further validate the quality and usefulness of the method developed.

ACKNOWLEDGMENTS

We thank Monika Destruelle (Qiagen) and Guido Meneghini (Perkin Elmer-Packard BioScience, Italy) for technical support. This work was supported by an AIRC-FIRC Research Fellowship (A.D.), a 2002 AIRC-FIRC grant (M.Z.), a Compagnia San Paolo Torino 2002 grant (M.Z.), a 2001 TIGEM-Telethon Regione Campania grant (M.Z.), and a PRIN 2002 grant MIUR-Rome (G.M. and M.Z.).

REFERENCES

1. **Kost, T.A. and J.P. Condreay.** 1999. Recombinant baculoviruses as expression vectors for insect and mammalian cells. Curr. Opin. Biotechnol. *10*:428-433.
2. **Hartl, F.U. and M. Hayer-Hartl.** 2002. Molecular chaperones in the cytosol: from nascent chain to folded protein. Science *295*:1852-1858.
3. **Yokoyama, N., M. Hirata, K. Ohtsuka, Y. Nishiyama, K. Fujii, M. Fujita, K. Kuzushima, T. Kiyono, and T. Tsurumi.** 2000. Co-expression of human chaperone Hsp70 and Hsdj or Hsp40 co-factor increases solubility of overexpressed target proteins in insect cells. Biochim. Biophys. Acta *1493*:119-124.2.
4. **Engel, M., O.G. Issinger, I. Lascu, T. Seib, S. Dooley, K.D. Zang, and C. Welter.** 1994. Phosphorylation of nm23/nucleoside diphosphate kinase by casein kinase 2 in vitro. Biochem. Biophys. Res. Commun. *199*:1041-1048.
5. **Freije, J.M., P. Blay, N.J. MacDonald, R.E. Manrow, and P.S. Steeg.** 1997. Site-directed mutation of Nm23-H1. Mutations lacking motility suppressive capacity upon transfection are deficient in histidine-dependent protein phosphotransferase pathways in vitro. J. Biol. Chem. *272*:5525-5532.
6. **Hartsough, M.T. and P.S. Steeg.** 2000. Nm23/nucleoside diphosphate kinase in human cancers. J. Bioenerg. Biomembr. *32*:301-308.
7. **Postel, E.H., V.H. Weiss, J. Beneken, and A. Kirtane.** 1996. Mutational analysis of NM23-H2/NDP kinase identifies the structural domains critical to recognition of a c-myc regulatory element. Proc. Natl. Acad. Sci. USA *93*:6892-6897.
8. **Lanio, T., A. Jeltsch, and A. Pingoud.** 2000. Automated purification of His6-tagged proteins allows exhaustive screening of libraries generated by random mutagenesis. BioTechniques *29*:338-342.
9. **Dell, A. and H.R. Morris.** 2001. Glycoprotein structure determination by mass spectrometry. Science *291*:2351-2356.
10. **Hoffmann, R., S. Metzger, B. Spengler, and L. Otvos, Jr.** 1999. Sequencing of peptides phosphorylated on serines and threonines by post-source decay in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. J. Mass Spectrom. *34*:1195-1204.
11. **Kuster, B. and M. Mann.** 1998. Identifying proteins and post-translational modifications by mass spectrometry. Curr. Opin. Struct. Biol. *8*:393-400.
12. **Zhou, H., J.D. Watts, and R. Aebersold.** 2001. A systematic approach to the analysis of protein phosphorylation. Nat. Biotechnol. *19*:375-378.
13. **Timmons, L., J. Xu, G. Hersperger, X.F. Deng, and A. Shearn.** 1995. Point mutations in awdKpn which revert the prune/Killer of prune lethal interaction affect conserved residues that are involved in nucleoside diphosphate kinase substrate binding and catalysis. J. Biol. Chem. *270*:23021-23030.
14. **Hamby, C.V., R. Abbi, N. Prasad, C. Stauffer, J. Thomson, C.E. Mendola, V. Sidorov, and J.M. Backer.** 2000. Expression of a catalytically inactive H118Y mutant of nm23-H2 suppresses the metastatic potential of line IV Cl 1 human melanoma cells. Int. J. Cancer *88*:547-553.

Received 25 February 2003; accepted 27 March 2003.

Address correspondence to Massimo Zollo, Telethon Institute of Genetics and Medicine, Via Pietro Castellino 111, 80131 Naples, Italy. e-mail: zollo@tigem.it

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 279, No. 36, Issue of September 3, pp. 37779–37788, 2004
Printed in U.S.A.

Phosphorylation by Protein Kinase CK2 Modulates the Activity of the ATP Binding Cassette A1 Transporter*

Received for publication, February 19, 2004, and in revised form, June 18, 2004
Published, JBC Papers in Press, June 24, 2004, DOI 10.1074/jbc.M401821200

Stein Roosbeek‡§, Frank Peelman‡, Annick Verhee‡, Christine Labeur‡, Hans Caster‡, Marc F. Lensink‡, Claudia Cirulli¶, Johan Grooten‖, Claude Cochet, Joël Vandekerckhove‡, Angela Amoresano¶, Giovanna Chimini‡‡, Jan Tavernier‡, and Maryvonne Rosseneu‡**

*From the ‡Department of Biochemistry, Faculty of Medicine and Health Sciences, Ghent University, B-9000 Ghent, Belgium, ‡‡Centre d'Immunologie F-13288 Marseille-Luminy, France, **Commissariat à l'Energie Atomique (Saclay, France), F-38054 Grenoble, France, ¶University of Naples, I-80126 Naples, Italy, and ‖Department of Molecular Biology, Ghent University, B-9000 Ghent, Belgium*

In a previous characterization of the ABCA subfamily of the ATP-binding cassette (ABC) transporters, we identified potential protein kinase 2 (CK2) phosphorylation sites, which are conserved in eukaryotic and prokaryotic members of the ABCA transporters (Peelman, F., Labeur, C., Vanloo, B., Roosbeek, S., Devaud, C., Duverger, N., Denefle, P., Rosier, M., Vandekerckhove, J., and Rosseneu, M. (2003) *J. Mol. Biol.* 325, 259–274). These phosphorylation residues are located in the conserved cytoplamic R1 and R2 domains, downstream of the nucleotide binding domains NBD1 and NBD2. To study the possible regulation of the ABCA1 transporter by CK2, we expressed the recombinant cytoplasmic domains of ABCA1, NBD1+R1 and NBD2+R2. We demonstrated that *in vitro* ABCA1 NBD1+R1, and not NBD2+R2, is phosphorylated by CK2, and we identified Thr-1242, Thr-1243, and Ser-1255 as the phosphorylated residues in the R1 domain by mass spectrometry. We further investigated the functional significance of the threonine and serine phosphorylation sites in NBD1 by site-directed mutagenesis of the entire ABCA1 followed by transfection into Hek-293 Tet-Off cells. The ABCA1 flippase activity, apolipoprotein AI and AII binding, and cellular phospholipid and cholesterol efflux were enhanced by mutations preventing CK2 phosphorylation of the threonine and serine residues. This was confirmed by the effect of specific protein kinase CK2 inhibitors upon the activity of wild type and mutant ABCA1 in transfected Hek-293 Tet-Off cells. The activities of the mutants mimicking threonine phosphorylation were close to that of wild type ABCA1. Our data, therefore, suggest that besides protein kinase A and C, protein kinase CK2 might play an important role *in vivo* in regulating the function and transport activity of ABCA1 and possibly of other members of the ABCA subfamily.

The role of the ABCA1[1] transporter, a member of the subfamily A of the ATP binding cassette transporters, in the efflux of cellular phospholipids and cholesterol has become well established (2). Several studies link ABCA1 mutations to impaired cellular cholesterol and phospholipid efflux characteristic of Tangier disease and high density lipoprotein-deficiency patients (3–5). Expression of WT and mutant ABCA1 in cultured cells demonstrated the correlation between the level of expression and activity of the ABCA1 transporter, the extent of binding to the apolipoprotein AI (apoAI) acceptor and the efflux of cellular lipid (6).

Human ABC transporters consist of a cytoplasmic nucleotide binding domain (NBD), which binds and hydrolyzes ATP, and of a membrane-spanning domain through which the substrate is translocated (7–8). Besides these elements, regulatory domains with putative phosphorylation sites were described in several human transporters (7). We carried out an extensive analysis of the subfamily A of the ABC transporters and showed that this subfamily consists of 13 human ABCA and of many eukaryotic and prokaryotic ABCA homologues (1). Multiple alignments of the subfamily A transporters demonstrated the high degree of sequence conservation and the particular topology of this subfamily. In the ABCA subfamily, the intracellular N-terminal residues are followed by a first transmembrane helix, by a long extracellular loop, and by a downstream cluster of five transmembrane helices separated by short loops. The transmembrane domain is followed by a cytoplasmic nucleotide binding domain NBD1 and by a downstream sequence of 80 conserved residues. In the full ABCA transporters, which have a high internal symmetry, these elements are repeated in the second half of the transporter (1). Among the human transporters of the ABCA subfamily, mutations in ABCA1 have been linked to Tangier disease, whereas mutations in ABCA2 impair the transport of retinyldiene and cause retinal degeneration (9). In both transporters, mutations in the nucleotide binding domains decrease the ATPase activity and thereby impair the transport activity of the mutants. Point mutations in the ABCA1 and ABCA2 extracellular loops probably affect the three-dimensional structure of the transporters and decrease their association with acceptor proteins required for substrate export (10).

The 80-residue segments downstream of the NBDs are conserved in the eukaryotic ABCAs and in the prokaryotic subfamily 7 precursors (1). In ABCA1 none of the natural mutations causing either Tangier disease or familial hypo-α-

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom correspondence should be addressed: University of Ghent, Department Biochemistry, Hospitaalstraat 13, B-9000 Ghent, Belgium. Tel.: 32-9-264-92-73; Fax: 32-9-264-94-96; E-mail: stein.roosbeek@UGent.be.

[1] The abbreviations used are: WT, wild type; ABCA1, ATP binding cassette A1; MALDI, matrix-assisted laser desorption ionization time; NBD, nucleotide binding domain; PKA and PKC, protein kinase A and C, respectively; CK2, protein kinase 2; Hek, human embryonic kidney; PS, phosphatidylserine; PE, phosphatidylethanolamine; GFP, green fluorescent protein; DMEM, Dulbecco's modified Eagle's medium; FBS, fetal bovine serum; PBS, phosphate-buffered saline.

lipoproteinemia occur in these sequences, although one of the Tangier mutations is located close to this domain (11). The functional significance of the conserved sequences downstream of the NBDs is still unknown; they might have a regulatory function, as proposed for other transporters (12). Phosphorylation by protein kinase A or C has been demonstrated for several ABC transporters, including cystic fibrosis transmembrane conductance regulator (13), P-glycoprotein (14), and recently ABCA1 (15–17). According to See *et al.* (15) serines 1042 and 2054 in NBD are phosphorylated by PKA *in vitro* and may potentially regulate ABCA1 function (15). Yamauchi *et al.* (16) report that apoAI activates PKC, resulting in phosphorylation and stabilization of ABCA1 (16). According to Martinez *et al.* (17), the potential phosphorylation sites for protein kinase 2 (CK2) and PKA in the ABCA1-PEST sequence, threonine 1286 and 1305, respectively, are constitutively phosphorylated and protect ABCA1 against calpain-mediated degradation. PKA and CK2 inhibitors did not, however, alter phosphorylation of these residues.

Among the kinases, protein kinase 2 or CK2 is a constitutively active Ser/Thr protein kinase essential for cell viability (18). CK2 can target at least 200 proteins (19) and is implicated in a wide variety of cellular functions such as tRNA and rRNA synthesis, apoptosis. and cell survival and transformation (20). The CK2 phosphoacceptor sites consist of clusters of acidic residues located mostly downstream from phosphorylatable seryl or threonyl residues. Prediction of consensus sequence motives in human ABCA1 using PROSITE identified 39 potential CK2 phosphorylation sites in ABCA1. Among these, the ^{1242}TTLEE and ^{2211}TTLDQ motives within the 80-residue R1 and R2 domains, downstream of, respectively, NBD1 and NBD2, were conserved in most members of the human ABCA subfamily (1). Another potential CK2 phosphorylation site, ^{1255}SGVD, is conserved in ABCA4, ABCA6–10, and ABCA12 but is absent in all NBD2s. Thus, it is conceivable that modulation of the ABCA1 transport activity might occur through serine or threonine phosphorylation by protein kinase CK2.

In this report we demonstrate that *in vitro* the NBD1+R1 and not the NBD2+R2 domain of ABCA1 is phosphorylated by CK2, and we used mass spectrometric methods to directly identify Thr-1242, Thr-1243, and Ser-1255 as the phosphorylated residues. We further investigated the functional significance of these CK2 phosphorylation sites in NBD1+R1 by site-directed mutagenesis of ABCA1. The ABCA1 flippase activity, the transporter ability to bind apolipoproteins AI and AII and to induce cellular phospholipid and cholesterol efflux, were enhanced by mutations preventing phosphorylation at that site. The specific CK2 inhibitors 4,5,6,7-tetrabromobenzotriazole and apigenin also enhanced the activity of ABCA1 transfected into Hek-293 Tet-Off cells, thus confirming the effect of the mutations. Our data, therefore, suggest that besides PKA and PKC, protein kinase CK2 might play an important role *in vivo* in regulating the function and transport activity of ABCA1 and possibly other members of the ABCA subfamily.

EXPERIMENTAL PROCEDURES

Materials—Tosylphenylalanyl chloromethyl ketone-treated trypsin, barium hydroxide, and α-cyano-4-hydroxycinnamic acid were from Sigma. Other mass spectrometry reagents and solvents were of highest purity from Carlo Erba (Milan, Italy). NBD-labeled cholesterol, phosphatidylserine (PS), phosphatidylethanolamine (PE), and phosphatidylcholine were purchased from Aventi Polar Lipids (Alabaster). Rhodamine hydrochloride and concanavalin A Alexa Fluor 633 were obtained from Molecular Probes (Leiden, The Netherlands). Doxycycline was purchased from Sigma. [*methyl*-^{3}H]Choline and [7-^{3}H]cholesterol were purchased from Amersham Biosciences. Hek-293 Tet-Off cells were purchased from Clontech (Erembodegem, Belgium). Recombinant apoAI was expressed in *Escherichia coli* (21), and plasma apoAII was prepared by ultracentrifugal isolation of high density lipoprotein, delipidation, and purification by DEAE chromatography as previously described (22). The CK2 inhibitor 4,5,6,7-tetrabromobenzotriazole was a gift of Dr. Shugar (Poland), Apigenin in Me_2SO was purchased from Sigma.

Construction of the Expression Vectors—For expression of the His-NBD1+R1 domain between residues 861 and 1306 and of His-NBD2+R2 between residues 1877 and 2232, vectors were constructed using the pET28a plasmid (Novagen, Madison, WI). DNA fragments encoding the protein sequences were PCR-amplified and cloned into the pET28a vector, and mutant plasmids were made by overlapping PCR. WT and mutant plasmids were checked by sequencing and by restriction analysis.

A bidirectional pBI-ABCA1/GFP vector encoding a hemagglutinin-tagged ABCA1/GFP fusion protein (23) was used to generate eight constructs encoding the threonine and serine ABCA1 mutants. The SplI/SplI fragment (nucleotides 3626–4535) was inserted into the SplI site of a pMCS5 vector (Molecular Biologische Technologie). Mutagenesis was performed by standard PCR, and mutated SplI fragments were transferred back into the SplI-digested pBI-ABCA1/GFP vector. All constructs were verified by restriction analysis and sequencing.

Expression and Purification of Soluble Recombinant His-NBD1+R1 and His-NBD2+R2 in E. coli—The length of the NBD1+R1 and NBD2+R2 constructs were selected to obtain soluble recombinant proteins and to avoid expression in the *E. coli* inclusion bodies (24). Proteins were expressed in transformed BL21(DE3)pLysS *E. coli* grown in LB medium (Invitrogen) under isopropyl-1-thio-β-D-galactopyranoside induction. Cells were centrifuged at 9000 × *g* for 10 min, washed with PBS, resuspended in 10 ml of lysis buffer (50 mM Tris-HCl, pH 7.6, 300 mM NaCl, 1 mM phenylmethylsulfonyl fluoride, 2 μg/ml leupeptin, 2 μg/ml aprotinin), and sonicated on ice 3 × 5 min at 60 watts.

Both His-NBD1+R1 and His-NBD2+R2 were purified by affinity chromatography on a Ni-PROBOND™ column. The column was equilibrated in a 50 mM Tris-HCl, pH 7.6, buffer containing 300 mM NaCl, 1 mM β-mercaptoethanol, 5% glycerol (v:v), 10 mM imidazole and incubated overnight at 4 °C with the cell lysate. Stepwise elution with 100 mM imidazole eliminated *E. coli* protein contaminants, whereas the His-NBD+R proteins eluted at 500 mM imidazole. The eluate was desalted on a HiTrap-Sepharose column and stored either at 4 °C or −80 °C. Sample purity was checked by 12% SDS-PAGE.

In Vitro Phosphorylation Assays—Recombinant CK2 and its α- and β-subunits were expressed in Sf9 cells and purified to homogeneity (25). For electrophoretic experiments, 2 or 8 μg of recombinant NBD1+R1, NBD2+R2, or mutants were incubated at 22 °C with either [γ-^{32}P]ATP or 10 μM ATP and 10 mM $MgCl_2$ in the presence of 0.3 μg of oligomeric CK2 or 1.0 μg of CK2α. The phosphorylated proteins were analyzed by SDS gel electrophoresis and autoradiography after 15 min. For mass spectrometric analysis, recombinant NBD+R proteins were incubated with CK2 for 30 min at 30 °C in the presence of 0.2 mM ATP. The reaction was terminated by 5 min of boiling and drying the samples.

Mass Spectrometric Analysis of the Phosphopeptides—Purified NBD+R proteins before and after treatment with CK2 were subjected to reduction with an excess of dithiothreitol, alkylation using iodoacetamide, and digestion with trypsin. To obtain the corresponding alkene moiety from the phosphate esters, the β-elimination reaction was optimized and simplified by incubating the peptide mixture in 5 M $Ba(OH)_2$ at 37 °C for 90 min under nitrogen. Solid carbonic dioxide was then added to eliminate the precipitated barium carbonate, which was removed by centrifugation at 13,000 × *g* for 5 min (26).

MALDI mass spectra were recorded using a Applied Biosystem Voyager DE-PRO instrument operating in reflector mode. A mixture of the analyte and of α-cyanohydroxycinnamic acid (10 mg/ml) in acetonitrile, ethyl alcohol, 0.1% trifluoroacetic acid (1:1:1, v/v/v) was applied to the metal sample plate and dried under vacuum. Mass calibration was performed with insulin at 5734.5 Da and a matrix peak at 379.3 Da as internal standards. Raw data were analyzed by using computer software provided by the manufacturer and reported as average masses.

Cell Cultures—Hek-293 Tet-off cells were cultured in DMEM media plus 10% fetal bovine serum (FBS), L-glutamine, and Geneticin on 12-well plates coated with poly-D-lysine. Cells were transfected for 24 h with the GFP-ABCA1 constructs using a mixture of polyethyleneimine and DNA at a 2/1 polyethyleneimine/DNA charge ratio and subsequently grown for 24 h. Cells were washed twice with phosphate-buffered saline (PBS) and incubated for 1 min with 500 μl of trypsin (Invitrogen). After the addition of 1 ml of PBS, 1% FBS and 0.5 mM EDTA, cells were filtered and kept on ice for flow cytometry. The percentage of transfected cells was determined by fluorescence-activated cell sorter

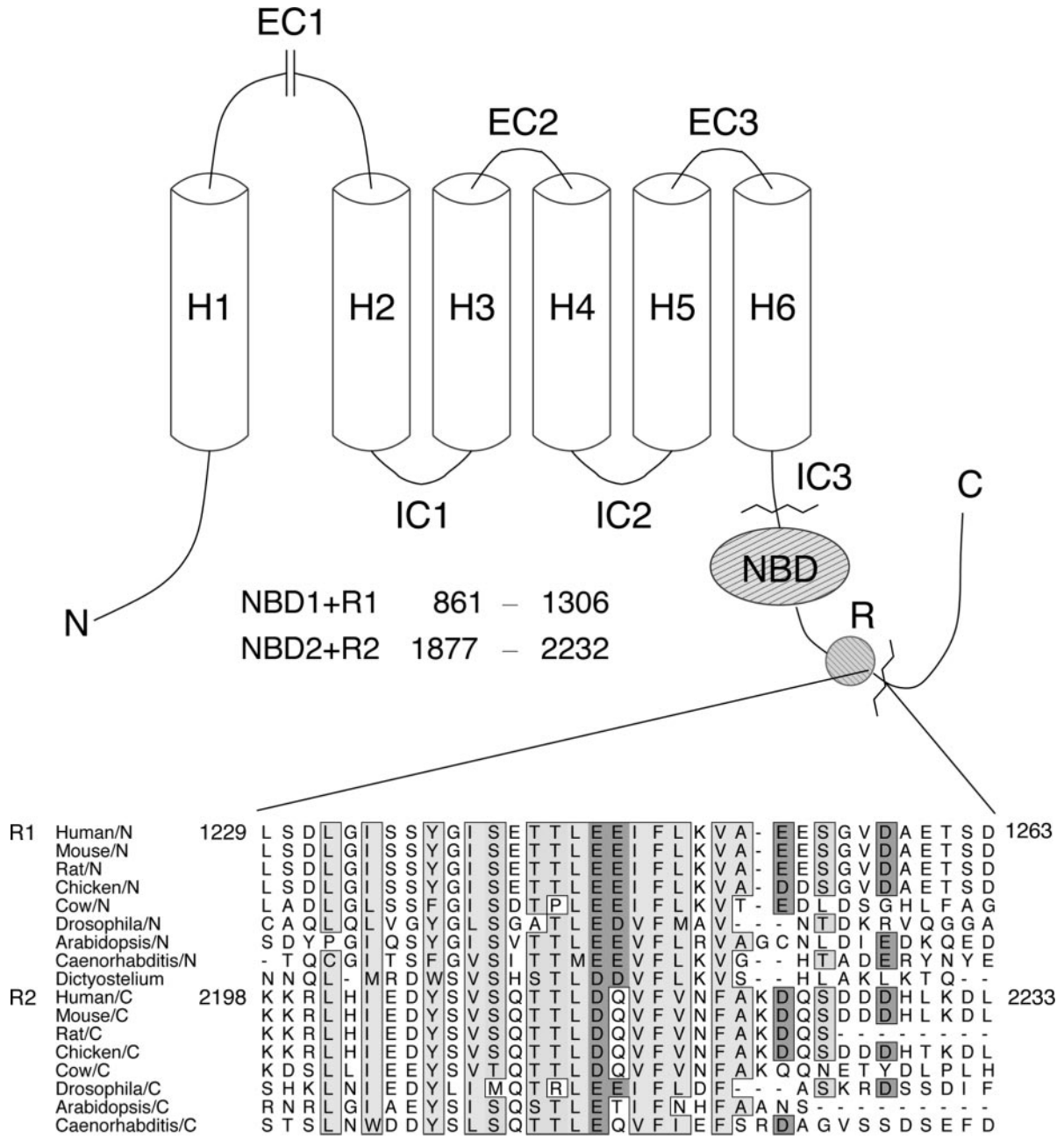


FIG. 1. **Topology of the N-terminal half of an ABCA full-transporter.** Six transmembrane helices are followed by the cytoplasmic NBD and the conserved regulatory R domain. The C-terminal half has an identical topology. The alignment of the R1 and R2 domains showing conservation of the potential CK2 phosphorylation sites at residues Thr-1242, Thr-1243 and Ser-1255 are shown for several species of ABCA1.

analysis on a FACScalibur (BD Biosciences) with an excitation wavelength of 488 nm and an emission wavelength of 530 nm.

Measurement of the ABCA1 Flippase Activity—Mock- and ABCA1-transfected Hek-293 Tet-Off cells were incubated at 37 °C for 45 min with 5 μM NBD lipid, washed with PBS at 4 °C, and incubated with 64 μM sodium dithionite in a Tris-HCl buffer, pH 10, for 5 min at 4 °C. The fluorescence intensity of the NBD probe before (I_{max}) and after (I_q) dithionite quenching was measured on a Bio-Tek FLX 800 microplate reader with excitation and emission wavelengths of 485 and 528 nm, respectively. Percent quenching was calculated as $(I_{max} - I_q)/(I_{max}) \times 100$.

Similar measurements were performed in the presence of 50 μM 4,5,6,7-tetrabromobenzotriazole and apigenin CK2 inhibitors on Mock-transfected and Hek-293 cells transfected with WT ABCA1 and with the T1242A, T1243A, T1242A/T1243A, and T1242D/T1243D ABCA1 mutants.

Monitoring of ApoAI and -AII Binding to Hek-293 Tet-Off Cells—After transfection with WT ABCA1, cells were washed twice with a 10 mM HEPES buffer, pH 7.4, containing 1.8 mM $CaCl_2$, 1 mM $MgCl_2$, 5 mM KCl, 150 mM NaCl and incubated with carboxyrhodamine-labeled recombinant apoAI or -AII at concentrations between 0 and 50 μg/ml for 1 h at 4 °C to monitor extent of binding. In subsequent experiments cells transfected with WT and mutant ABCA1 were incubated with 20 μg/ml apoAI or -AII for 1 h at 4 °C. Cells were detached by mild trypsinization for 2 min at 37 °C using 0.05 g/liter trypsin in PBS and kept on ice, and flow cytometric recordings were performed. Cells were subdivided into three populations as a function of their level of ABCAI expression, reflected by the GFP fluorescence intensity. Apoprotein binding data were calculated for each cell population based upon the mean carboxyrhodamine fluorescence intensity (27).

Phospholipid and Cholesterol Efflux to ApoAI and ApoAII Induced by ABCA1—Phospholipid and cholesterol efflux were measured using both radioactivity and fluorescence. For radioactive measurements, Hek-293 Tet-Off cells transfected with WT and mutant ABCA1 were washed twice with DMEM and incubated either with 1 μCi/ml [*methyl*-^{3}H]choline or [^{3}H]cholesterol in DMEM plus 2.5% FBS. After 24 h cells were washed twice with DMEM and incubated with 1 ml of DMEM plus 0.2% bovine serum albumin for 16 h at 37 °C. Cells were washed twice with DMEM and incubated at 37 °C with 20 μg/ml recombinant apoAI for 4 h

for [*methyl*-^{3}H]choline efflux or 6 h for [^{3}H]cholesterol efflux. Supernatant was collected and centrifuged for 10 min at 810 × *g*. 500 μl of supernatant was transferred into 5 ml of scintillation liquid (Optiphase Hisafe 3). Cells were washed with PBS and lysed with 1 ml of NaOH (0.1 N). 500 μl of cell suspension was transferred into 5 ml of scintillation liquid (Optiphase Hisafe 3). ^{3}H counts per minute (cpm) were measured in a liquid scintillation counter (Wallac 1409), and % lipid efflux was calculated as (cpm in supernatant/(cpm in supernatant + cpm in cells)) × 100. For measurements using fluorescence, Hek-293 Tet-Off cells transfected with WT and mutant ABC-A1 were labeled with either 5 μM NBD phospholipid or 2.5 μM NBD cholesterol for 45 min at 4 °C. Cells were washed 3 times with 1 ml of PBS and incubated with 500 μl of DMEM containing either 20 μg/ml recombinant apoAI or 40 μg/ml apoAII. Efflux was measured either after 4 h of incubation at 37 °C for NBD phospholipids or after 6 h for NBD cholesterol by measuring NBD fluorescence intensity. Percentage efflux was calculated as $((I_{control} - I_{apoAI})/I_{control}) \times 100$.

Confocal Microscopy—Cellular localization of the expressed ABCA1 WT and mutants was analyzed by confocal microscopy using GFP fluorescence at excitation and emission wavelengths of, respectively, 488 and 509 nm. Hek-293 Tet-Off cells were plated on poly-D-lysine-coated Lab-Tek® German borosilicate confocal microscope slides at a density of 2.5 × 10^4 cells/chamber in 400 μl in DMEM plus 10% FBS, transfected, grown for 48 h in DMEM plus 10% FBS and washed twice with PBS. Cells were incubated with 25 μg/ml concanavalin A Alexa fluor 633, an endoplasmic reticulum-specific fluorescence dye, for 30 min. Cells were washed with PBS, and the Alexa fluor 633 fluorescence was measured using the excitation and emission wavelengths of, respectively, 632 and 647 nm.

ABCA1 Half-life Measurement—To determine the half-life of WT and mutant ABCA1, the transporter flippase activity was measured as a function of time in transfected Hek-293 Tet-Off cells, which were treated with 100 ng/ml doxycycline to stop transcription of the ABCA1 constructs (28, 29). The ABCA1 flippase activity for phosphatidylethanolamine was measured as described above using NBD PE after 0, 30, 60, 90, and 120 min. The half-life was estimated from half the total activity decay (28).

RESULTS

Identification of Protein Kinase CK2 Phosphorylation Sites on ABCA1—Fig. 1 shows a schematic diagram of the N-terminal half of ABCA1, consisting of a membrane-spanning domain made of six helices followed by the cytoplasmic NBD and the regulatory R domain. The C-terminal half of ABCA1 has an identical topology (1). Among the 39 potential CK2 phosphorylation sites predicted by PROSITE in ABCA1, we focused upon those located in NBD1, NBD2, and in the conserved downstream R1 and R2 domains (1). Eight potential CK2 sites are predicted within the cytoplasmic NBD1+R1 domain between residues 861 and 1306. One site at ^{989}TVEE is in NBD1, whereas three sites at ^{1242}TTLE, ^{1243}TLEE, and ^{1255}SGVD are located in the R1 domain. By comparison, seven CK2 sites are predicted between the C-terminal residues 1877–2232, four of which are in NBD2 and two, ^{2211}TTLD and ^{2225}SDDD, are in the R2 domain. We examined the conservation of the human ABCA1 Thr-1242, Thr-1243, and Ser-1255 residues in ABCA sequences of other species (Fig. 1). As previously reported for other transporters of the ABCA family (1), the threonine residues are better conserved in the N-terminal R1 compared with the R2 domain, whereas Ser-1255 is conserved only in mammalian ABCA.

To investigate the *in vitro* phosphorylation properties of the ABCA1 cytoplasmic domains, we expressed the NBD+R domains in *E. coli* and used the purified recombinant proteins as substrates for protein kinase CK2 assays *in vitro* using both the CK2 holoenzyme or its isolated α catalytic subunit. Fig. 2 shows that only ABCA1 NBD1+R1 and not NBD2+R2 is phosphorylated *in vitro* by CK2. Interestingly, only the CK2 holoenzyme but not its α-subunit was able to phosphorylate the NBD1+R1 substrate, indicating that the presence of the CK2β regulatory subunit is required for efficient phosphorylation.

As the ^{1242}TTLE, ^{1243}TLEE, and ^{1255}SGVD are the best

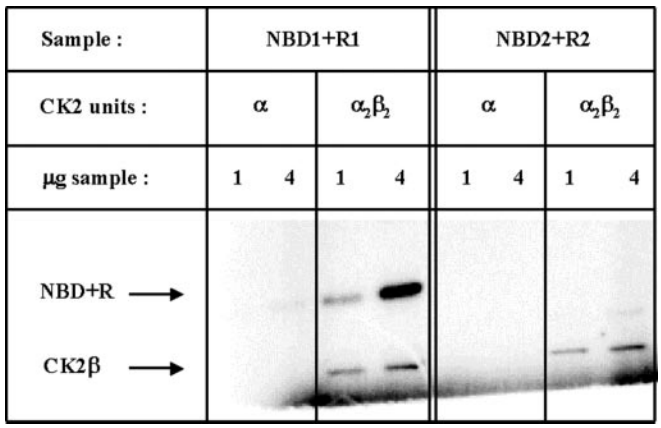


FIG. 2. ***In vitro* phosphorylation of recombinant NBD1+R1 (residues 861–1306) and NBD2+R2 (residues 1877–2232) proteins by the CK2 holoenzyme (α2β2) and by its α catalytic subunit.** Either 1 or 4 μg recombinant NBD+R were incubated with 0.3 μg of CK2$\alpha_2\beta_2$ or 1.0 μg of CK2α and P^{32}-labeled ATP. Phosphorylated proteins were detected by autoradiography.

conserved potential CK2 phosphorylation sites in NBD1, we carried out site-directed mutagenesis at these positions both in the recombinant NBD1+R1 protein and in the entire ABCA1 construct. In both systems we engineered several mutants designed at preventing phosphorylation by mutation either of the target threonines and serine, T1242A, T1243A, T1242A/T1243A, S1255A, or of the downstream cluster of acidic residues, E1245Q, E1246Q, E1245Q/E1246Q. Mutants were also designed to mimic phosphorylated residues as T1242D, T1243D, T1242D/T1243D.

All NBD1+R1 mutants were expressed as soluble proteins in *E. coli* and purified at yields comparable with WT NBD1+R1 (24). Using recombinant CK2, phosphorylation of the Thr-Ala mutants was not completely abolished, suggesting the presence of multiple CK2 phosphorylation sites within this domain (data not shown).

This was confirmed by MALDI mass spectrometric analysis of the native and phosphorylated recombinant NBD+R proteins (Table I). The phosphopeptides were identified by their mass differences due to the presence of phosphate moieties compared with the peptide masses expected from the protein sequence. In agreement with the phosphorylation assays, we observed no phosphorylation of NBD2+R2; in WT NBD1+R1 we observed that the signal at *m/z* 4542.5 corresponds to the ABCA1 peptide 1229–1269, carrying three phosphate groups at Thr-1242, Thr-1243, and Ser-1255. The peptide mixture was then submitted to the β-elimination reaction followed by MALDI mass spectrometric analysis. Under strong alkaline conditions elimination of the phosphate moiety on Ser(P) and Thr(P) to form dehydroalanine (ΔSer) or dehydroalanine-2-butyric acid (ΔThr) yields new signals corresponding to the β-eliminated peptides. The signal at *m/z* 4249.5 was assigned to the 1229–1269 ABCA1 peptide, occurring 299 Da lower, thus confirming phosphorylation at the three positions. Analysis of the T1242A, T1243A, and T1242A/T1243A NBD1 mutants supports the loss of either one or two threonine phosphorylation sites in the mutated peptides, whereas the Ser-1255 site is preserved. Analysis of the S1255A mutant demonstrates the loss of one phosphorylation site in the Val-1251—Arg-1272 tryptic peptide (Table I). These data confirm that besides Thr-1242 and Thr-1243, Ser-1255 is another CK2 phosphorylation site in ABCA1. None of the predicted CK2 phosphorylation sites in the NBD1 (residues 866–995) or between NBD1 and the R1 domain was phosphorylated.

ABCA1 Activity Increases Linearly with Protein Expression—Hek-293 Tet-Off cells were transfected with varying amounts

TABLE I
Mass spectrometric analysis of NBD1+R1 phosphorylation

NBD1+R1	Tryptic peptide	Molecular mass	Molecular mass after β-elimination	Phosphorylated residues
		Da	*Da*	
WT	Leu-1229–Arg-1269	4542.5	4249.5	Thr-1242, Thr-1243, Ser-1255
T1242A	Leu-1229–Lys-1250	2483.6	2385.9	Thr-1243
T1242A	Val-1251–Arg-1272	2411.6	2313.7	Ser-1255
T1243A	Leu-1229–Lys-1250	2483.1	2385.2	Thr-1242
T1243A	Val-1251–Arg-1272	2411.2	2313.2	Ser-1255
T1242A/T1243A	Leu-1229–Lys-1250	2356.1	2356.1	
T1242A/T1243A	Val-1251–Arg-1272	2411.6	2313.7	Ser-1255
S1255A	Val-1251–Arg-1272	2297.7	2297.7	

of ABCA1-GFP DNA, between 1 and 10 μg. Extent of transfection was estimated as the product of the GFP fluorescence intensity and the number of transfected cells, both measured by fluorescence-activated cell sorter analysis (23). As a result of an increased flippase activity, the percentage of NBD PE in the plasma membrane outer leaflet increased linearly between 20 and 58% for a 5-fold increase of WT ABCA1 transfection (Fig. 3*A*). Parallel data were obtained for both apoAI binding to transfected cells (Fig. 3*B*) and phospholipid efflux (Fig. 3*C*). This linear relationship between ABCA1 expression and activity was used to normalize the activity of ABCA1 mutants compared with WT ABCA1.

Mutation of the CK2 Phosphorylation Sites Increases ABCA1 Flippase Activity—To assess the effect of CK2 phosphorylation on the ABCA1 flippase activity we measured phospholipid flip-flop by WT ABCA1 and the mutants. Incubation of ABCA1-transfected Hek-293 Tet-Off cells with NBD phospholipids at 37 °C and subsequent quenching of the fluorescence with sodium dithionite at 4 °C yielded the percentage of labeled phospholipid in the outer leaflet of the plasma membrane. In mock-transfected cells there was around 30 and 25% NBD PE and NBD PS, respectively, in the plasma membrane outer leaflet. After transfection with 3 μg of ABCA1-GFP DNA, these percentages increased to 38 and 35%, indicative of phospholipid flip-flop by ABCA1.

Compared with WT ABCA1, set as 100%, the PE and PS content in the outer leaflet of the plasma membrane increased up to 130 ± 14 and 123 ± 8% for the T1242A mutant, 154 ± 15 and 148 ± 11% for the T1243A mutant, 170 ± 15 and 207 ± 16% for the T1242A/T1243A mutant, 141 ± 11 and 136 ± 10% for the S1255A mutant, and 141 ± 13 and 120 ± 9% for the E1245Q/E1246Q mutant, respectively (Fig. 4). All measurements were performed in triplicate, and the differences were significant at $p < 0.01$ using a Student t test. In contrast, when the threonine residues were mutated to aspartic acid to mimic a phosphorylated residue, the phospholipid content in the membrane outer leaflet decreased compared with that of the Thr-Ala mutants and came closer to that of WT ABCA1 (Fig. 4).

The effect of the specific CK2 inhibitors 4,5,6,7-tetrabromobenzotriazole and apigenin on the flippase activity in Hek-293 Tet-Off cells transfected with WT ABCA1 and with the T1242A, T1243A, T1242A/T1243A, S1255A, and T1242D/T1243D mutants was consistent with the above data. Incubation with 50 μM either 4,5,6,7-tetrabromobenzotriazole or apigenin increased the flippase activity of WT ABCA1 by 30 ± 2.6 and 25 ± 3.4% ($n = 3$, $p < 0.001$), respectively, whereas the flippase activity increased only by around 5 ± 0.7% ($n = 3$, $p < 0.01$) in cells transfected with ABCA1 mutants.

Mutations of the ^{1242}TTLEE and ^{1255}SGVD CK2 Phosphorylation Sites Modulate the Binding of ApoAI and ApoAII to ABCA1—We compared the binding of carboxyrhodamine-labeled apoAI and apoAII to the Hek-293 Tet-Off cells transfected with WT and ABCA1 phosphorylation mutants by flow cytometric analysis (27). We analyzed apolipoprotein binding to three cell populations, with low, middle, and high levels of transfection, estimated from the GFP fluorescence, with cut-offs at >7 and <20 for population 1 (ABCA1+), >21 and <90 for population 2 (ABCA1++), and >91 for population 3 (ABCA1+++) (27). Subtraction of the carboxyrhodamine fluorescence intensity of Mock-transfected cells yielded saturation curves, with a plateau above 10 μg/ml for apoAI/Rho and apoAII/Rho binding to WT ABCA1 (Fig. 5). Apparent binding constants to the three cell populations were 7, 30, and 24 × 10^{-8} M for apoAI binding compared with 4, 16, and 17 × 10^{-8} M for apoAII binding. These values are comparable with those measured for apoAI/Cy5 binding to WT ABCA1 (27).

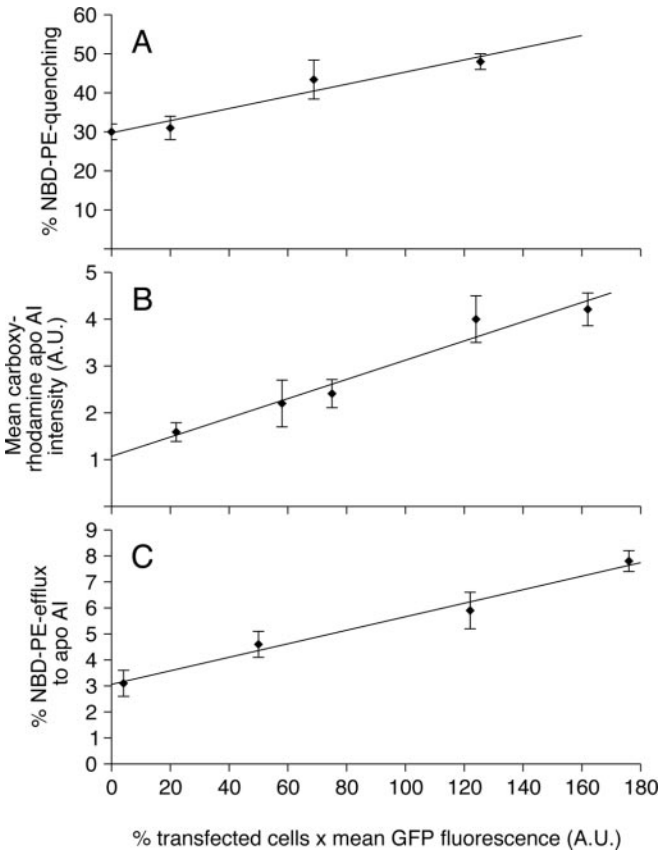


FIG. 3. **Relationship between WT ABCA1 activity and the extent of WT ABCA1 transfection into Hek-293 Tet-Off cells.** *A*, percentage of NBD PE quenching by sodium dithionite as a measure of the NBD PE flip-flop from the membrane inner to outer leaflet. *B*, specific binding of apoAI to ABCA1-transfected cells. *A.U.*, absorbance units. *C*, NBD PE efflux to apoAI.

The carboxyrhodamine fluorescence intensity due to specific binding of 20 μg/ml apoAI or apoAII to cells transfected with ABCA1 mutants was expressed relative to WT ABCA1 (Fig. 6). ApoAI- and apoAII-specific binding to the cells transfected with the single threonine and serine ABCA1 mutants increased

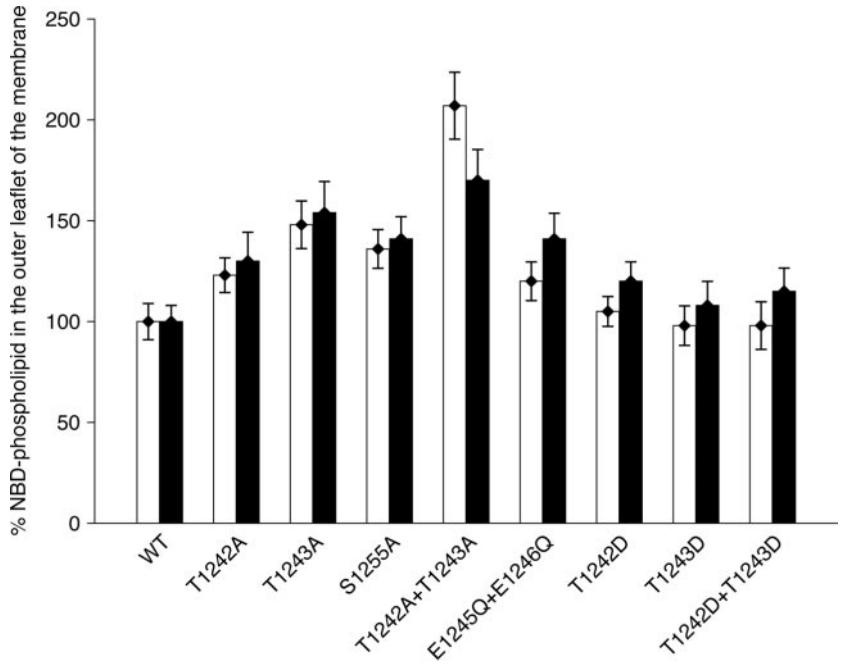


FIG. 4. **NBD phospholipid quenching in the plasma membrane of Hek-293 Tet-Off cells transfected with WT and mutant ABCA1.** Quenching by the ABCA1 mutants is expressed relative to WT ABCA1 after normalization. □, NBD PS; ■, NBD PE.

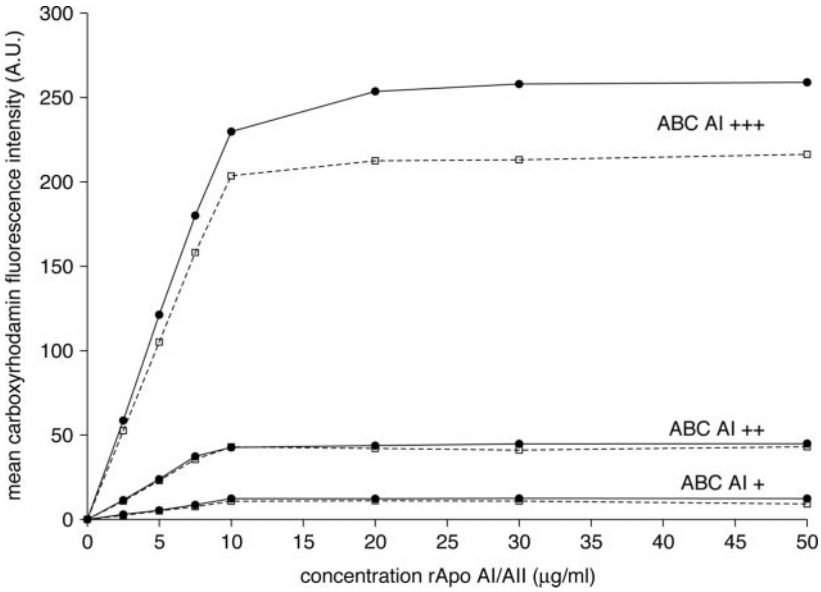


FIG. 5. **Fluorescence intensity of carboxyrhodamine-labeled-apoAI (□) and -apoAII (■) bound to Hek-293 Tet-Off cells transfected with WT ABCA1 as a function of the apoprotein concentration.** Binding was calculated for three cellular populations transfected at a high (*ABCA1*+++; ABCA1-GFP fluorescence >91), intermediary (*ABCA1*++; 91> ABCA1-GFP fluorescence >21), and low level (*ABCA1*+; 21 > ABCA1-GFP fluorescence >7) *A.U.*, absorbance units.

compared with WT ABCA1. This increase was more pronounced for the T1243A mutation either alone or in combination with the T1242A mutation, whereas the effect of the T1242A and S1255A mutants was similar. The E1245Q/E1246Q mutation of the downstream acidic residues, which also impairs CK2 activity, similarly increased cellular apoAI and apoAII binding. The T1242D, T1243D, T1242D/T1243D mutations, which mimic CK2 phosphorylation, had an opposite effect to the Thr-Ala mutations, as apoprotein binding to cells transfected with these mutants amounted to around 110% of WT ABCA1.

Mutation of the CK2 Phosphorylation Sites in ABCA1 Modulates Phospholipid and Cholesterol Efflux to ApoAI and ApoAII—Phospholipid and cholesterol efflux to apoAI and apoAII together with phospholipid distribution across the plasma membrane were measured for Hek-293 Tet-Off cells transfected with WT ABCA1 and with the phosphorylation mutants. Data were normalized to account for differences in expression levels between WT ABCA1 and mutants. ABCA1 lipid efflux was measured using both radioactive-labeled and fluorescent-labeled lipids, yielding comparable data. Only the results for NBD-labeled lipids are given in Fig. 7, *A* and *B*.

Incubation of the transfected Hek-293 Tet-Off cells, labeled with NBD phospholipids or NBD cholesterol with either unlabeled apoAI or apoAII decreased the NBD lipid fluorescence intensity due to the efflux of labeled lipids to the unlabeled apoprotein acceptors. The saturation curves for the percentage of phospholipid efflux were similar to the apoprotein binding curves and reached a plateau around 20 μg/ml apoAI or apoAII. The percentages of NBD-labeled phosphatidylethanolamine and cholesterol efflux to apoAI induced by WT ABCA1 were around 7 and 4%, respectively. Using radioactive-labeled PE and cholesterol, efflux percentages were, respectively, 8.5 and 4.7%, comparable with those measured using fluorescence. PE and PS efflux to apoAI and apoAII (Fig. 7*A*) and cholesterol efflux to apoAI (Fig. 7*B*) were expressed relative to WT ABCA1 for all mutants after normalization. The T1242A and T1243A mutations either alone or in combination and the E1245Q/E1246Q mutation, which all decrease CK2 activity, increased phospholipid and cholesterol efflux (Fig. 7, *A* and *B*). The effect

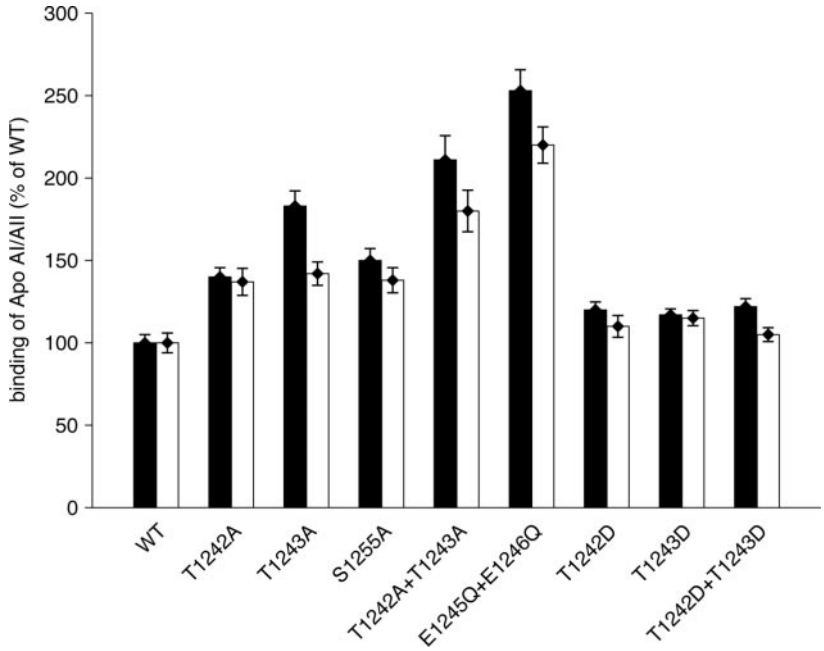


FIG. 6. **Carboxyrhodamine-labeled-apoAI (■) and -apoAII (□) binding to Hek-293 Tet-Off cells transfected with WT and mutant ABCA1.** The apoprotein binding to mutant transfected cells is expressed relative to WT ABCA1 after normalization.

of the S1255A mutant on lipid efflux was comparable with the single threonine mutants. The T1242D, T1243D, and T1242/T1243D mutations had an opposite effect as the efflux induced by that mutant was close to WT ABCA1. These effects parallel those observed for apoprotein binding to the mutant-transfected cells (Fig. 6).

Cellular Localization of the ABCA1 WT and Mutants—Using confocal microscopy, we observed the same cellular localization of WT and mutant ABCA1, which were mainly found at the plasma membrane of transfected Hek-293 Tet-Off cells. Incubation of the ABCA1-transfected cells with the endoplasmic reticulum-specific fluorescence dye concanavalin A Alexa fluor 633 showed that only a small percentage of ABCA1 WT and mutants is localized in the endoplasmic reticulum (Fig. 8). These results indicate that mutations of the CK2 phosphorylation sites have little effect upon the cellular trafficking of the ABCA1 transporter.

ABCA1 WT and Mutants Have Similar Half-lives—The decrease of ABCA1 flippase activity as a function of time was used to estimate the transporter half-life. In agreement with previous reports, we observed a short half-life of ~40 min for WT ABCA1 (30). All mutants had similar half-lives compared with wild type ABCA1, varying between 35 and 50 min (data not shown), indicating that mutations of the CK2 phosphorylation sites do not significantly affect the ABCA1 turnover.

DISCUSSION

In this paper we showed that besides PKA (15) and PKC (16–17), protein kinase CK2 is an important kinase to regulate ABCA1 activity. We further demonstrated that phosphorylation occurs at conserved threonine and serine residues in the R1 domain, downstream of NBD1. CK2-specific phosphorylation in the R1 domain affects the major functions of the ABCA1 transporter, including phospholipid flip-flop, ABCA1 binding to apoAI and apoAII, and cellular phospholipid and cholesterol efflux, in a coordinated manner. Mutations aimed at suppressing the phosphorylation sites increase these activities, whereas the activity of mutants mimicking phosphorylation of the threonine residues is close to that of WT ABCA1. Cellular treatment with specific CK2 inhibitors further supports these data.

Using recombinant WT and mutant NBD1+R1 proteins, we further showed that they are *in vitro* substrates for CK2. Combined mutation of either the two threonines or the downstream acidic residues in the recombinant NBD1+R1 domain did not completely suppress protein phosphorylation. This is in agreement with the mass spectrometric identification of the CK2 phosphorylation sites in NBD1+R1, identifying residue Ser-1255 besides Thr-1242 and Thr-1243 as a putative ABCA1 phosphorylation site. This was then confirmed by the engineering of the S1255A mutant, whose activities are comparable with those of the T1242A mutant. Threonines 1242 and 1243 had been identified as some of the best conserved residues belonging to potential CK2 phosphorylation sites in cytoplasmic ABCA1 domains. These residues are conserved in most of the human and mammalian ABCA transporters and in several other species and even in family 7 bacterial transporters (1). The serine 1255 residue is not as highly conserved as the threonines among the different ABCA1 species or within the ABCA family. However, because serine is often the preferred substrate of CK2 (20), it might constitute a primary phosphorylation site in ABCA1, in analogy with the data obtained for PTEN phosphatidylinositol 3′-phosphatase (31). The alignment of the R1 and R2 domains shows the correspondence between the consensus CK2 phosphorylation sequence around residue Thr-1242 in R1 and residue Thr-2211 in R2. Threonine 2212 in NBD2 does not belong to a CK2 consensus sequence, whereas Ser-1255 in R1 has no counterpart in R2. As shown both by autoradiography and mass spectrometry, CK2 phosphorylation seems restricted to the N-terminal R1 domain of ABCA1. Because mutations of the CK2 phosphorylation sites and treatment with specific CK2 inhibitor both increase the ABCA1 flippase and cellular efflux activities as well as apoprotein binding, CK2 phosphorylation of the ABCA1 R1 domain might contribute to the *in vivo* regulation of ABCA1 functionality.

Although there is a high degree of sequence homology between the N- and C-terminal cytoplasmic domains of ABCA1, several reports suggest that NBD1 and NBD2 might have specific functions. There are differences between the NBDs ATPase activity, as we measured a 3-fold higher activity for recombinant ABCA1 NBD1 compared with NBD2 (24). Folding and association of the two NBD domains are not identical (32–33), and several reports suggest that NBD1 and NBD2 share a common interface in ABC full transporters (32–33). The activation cycle of the ABCA transporters consists of ATP binding and hydrolysis, ADP dissociation accompanied by con-

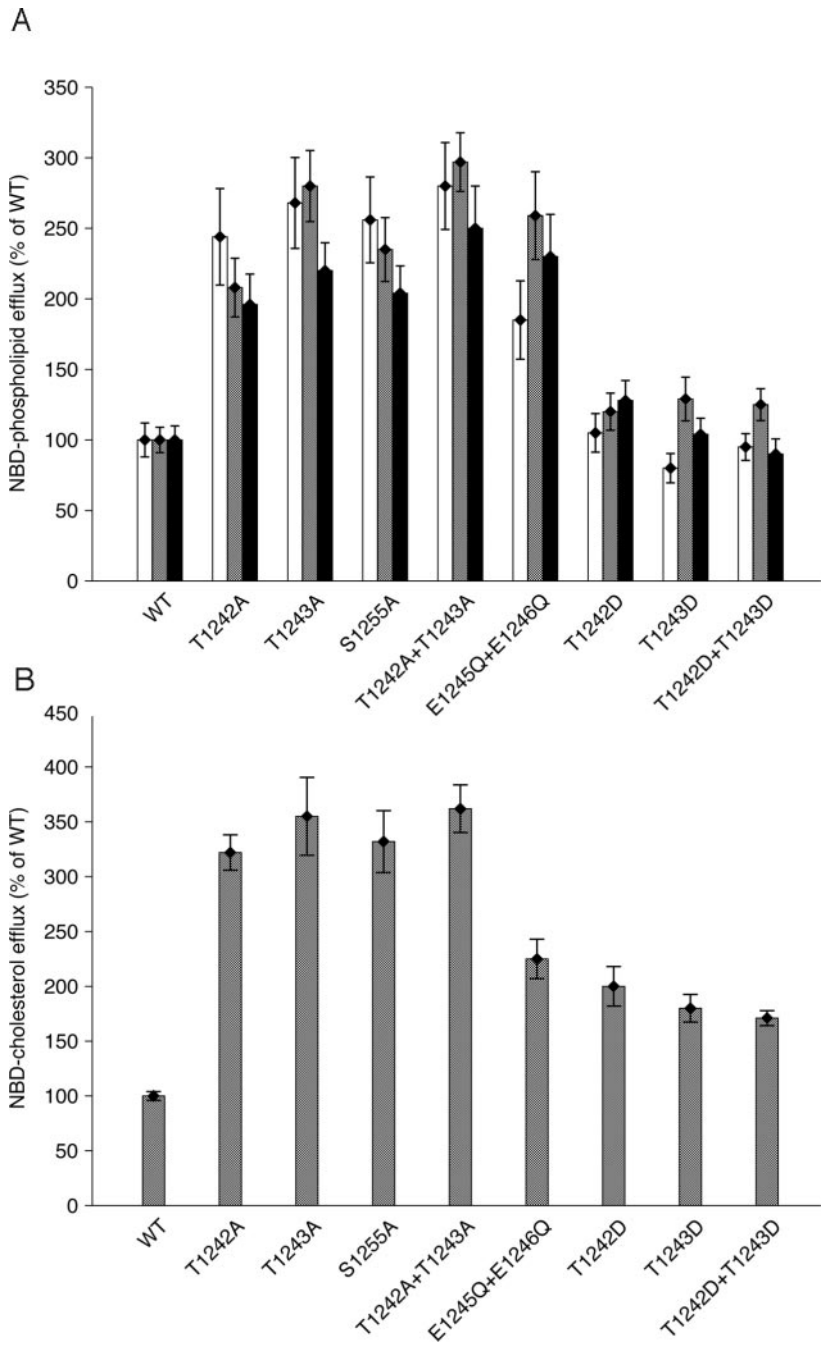


FIG. 7. *A*, NBD phospholipid efflux to apoAI and apoAII after incubation of transfected Hek-293 Tet-Off cells with 20 μg/ml apoAI or apoAII for 6 h at 37 °C. Efflux by ABCA1 mutants is expressed relative to WT ABCA1 after normalization. *Open squares*, NBD PS efflux to apoAI; *gray squares*, NBD PE efflux to apoAI; *filled squares* NBD PE efflux to apoAII. *B*, NBD cholesterol efflux to apoAI (*filled squares*) after incubation of the transfected cells with 20 μg/ml apoAI for 6 h at 37 °C. Efflux by ABCA1 mutants is expressed relative to WT ABCA1 after normalization.

formational changes in the NBD domains, and finally, of substrate transport across the membrane. This process is not random but is probably initiated through ATP binding to the nucleotide binding domain with highest affinity (34). It is quite conceivable that phosphorylation of the R1 domain by CK2 could bring a conformational change to this domain and consequently to NBD1 to start the activation cycle.

The recent x-ray structure of the vitamin B12 transporter BtuCD (33) suggests that the two NBD domains interact with each other in an anti-parallel way and that these interactions are stabilized by interactions between the NBD downstream domains. The NBD downstream domains in the maltose transporter correspond to the conserved R1 and R2 domains identified in the ABCA subfamily. Interactions between R1 and R2 in the ABCA subfamily might enhance the cooperativity between the N- and C-terminal half of the ABCA transporters and, thus, contribute to a fine regulation of the transporter activity. These interactions are likely to be dependent upon the phosphorylation state of the protein domain. Serine and threonine phosphorylation by a cytoplasmic CK2 kinase might, thus, represent another mechanism for the intracellular regulation of the transporter activity.

Protein kinase CK2 is a constitutively active kinase that is implicated in a variety of cellular functions including cell proliferation, differentiation, and survival (19). Our experiments showed that the regulatory β subunit of CK2 is required to achieve optimal phosphorylation of the ABCA1 substrate. Thus, this transporter belongs to the class I of the CK2 substrates whose phosphorylation is strictly dependent upon the presence of the CK2β subunit (35). On the other hand, we observed an increased functionality of the protein when its phosphorylation is abrogated, suggesting that CK2 acts as a down-regulator of ABCA1 function. A similar function of CK2 was previously described for Engrailed 2 secretion (36).

Regulation of ABCA1 might occur in different ways; phosphorylation might change the expression or cellular location of

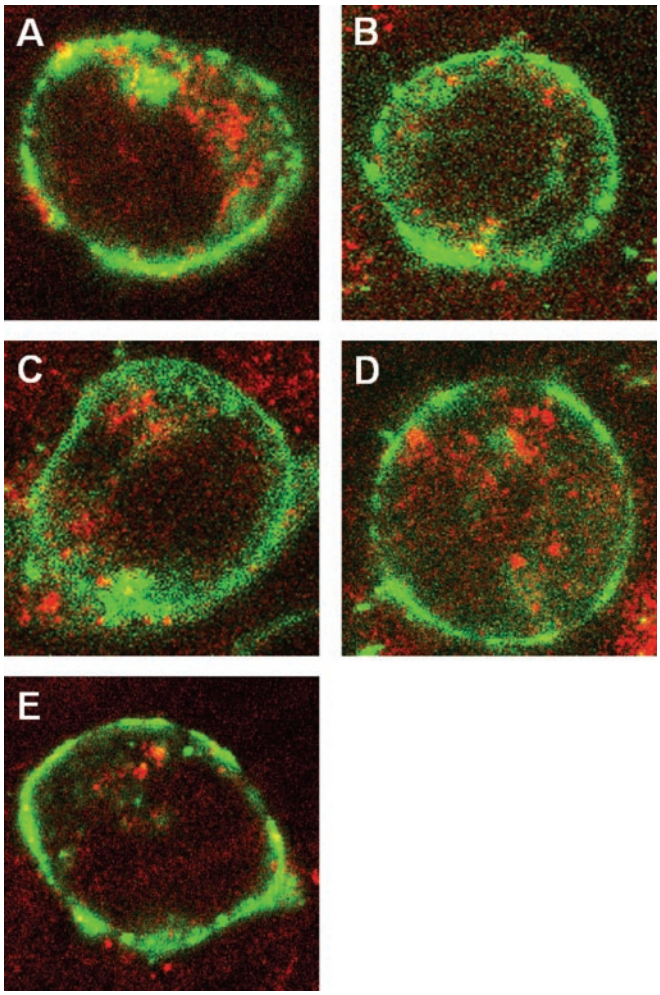


FIG. 8. **Cellular localization of GFP-tagged ABCA1 WT and mutants studied by confocal microscopy.** *Green fluorescence* corresponds to the GFP fluorescence, and *red fluorescence* corresponds to the endoplasmic reticulum-specific probe concanavalin A Alexa fluor 633. *Yellow fluorescence* indicates the colocalization of the GFP-tagged ABCA1 and the endoplasmic reticulum. *A*, ABCA1 WT; *B*, T1243A mutant; *C*, T1242A/T1243A mutant; *D*, T1242D/T1243D mutant; *E*, S1255A mutant.

ABCA1. This is, however, not likely as 1) our confocal microscopic data showed similar locations at the plasma membrane, and 2) the expression of the Thr-Ala and Ser-Ala mutants was lower than of WT ABCA1. We also showed that ATP binding and ATPase activity are not affected by the threonine to alanine or by the glutamic acid to glutamine mutations (23). The Trp exposure and the secondary structure of the various mutants of the NBD1+R1 domain were quite similar. The structural changes due to phosphorylation are, therefore, probably minimal, and a likely effect of CK2 phosphorylation is, therefore, the modulation of the cooperativity between NBD1 and NBD2 in native ABCA1.

The activities of several ABC transporters are regulated by their phosphorylation, as demonstrated for the cystic fibrosis transmembrane regulator, which is phosphorylated both by PKA and PKC (13). Phosphorylation of the P-glycoprotein by PKA and PKC seems to modulate swelling-activated Cl^- currents (14), not drug transport (37). PKA activation by cAMP analogues increased ABCA1 phosphorylation and cellular cholesterol efflux (15). Mutation of a PKA phosphorylation site in ABCA1 impaired apoAI-mediated release of cellular lipid (15). According to Yamauchi *et al.* (16) ABCA1 phosphorylation is, furthermore, critical for the stability and integrity of the transporter (16). PKC_α is activated by apoAI, and it subsequently phosphorylates and stabilizes ABCA1. According to Martinez *et al.* (17), a PEST sequence in ABCA1 is critical for apoAI protection against ABCA1 degradation by calpain. This sequence includes the casein kinase CK2 and PKA potential phosphorylation sites Thr-1286 and Thr-1305, respectively. These authors found little effect of the casein kinase inhibitors on ABCA1 phosphorylation, in contrast with our findings, and concluded that CK2 probably does not phosphorylate the Thr-1286 residue in the PEST sequence. Although this residue belongs to a potential CK2 phosphorylation site in ABCA1, it is localized in a low conservation sequence in the ABCA family and might not become phosphorylated, in contrast to the conserved Thr-1242, Thr-1243, and Ser-1255 residues tested in our study. Furthermore, we measured the effect of the CK2 inhibitors on the specific ABCA1 activity, whereas Martinez *et al.* (17) investigated the extent of total ABCA1 phosphorylation by autoradiography. Besides PKA and PKC, which increase ABCA1 activity, CK2 might play an opposite role and act as a down-regulator of ABCA1 in the cell.

REFERENCES

1. Peelman, F., Labeur, C., Vanloo, B., Roosbeek, S., Devaud, C., Duverger, N., Denefle, P., Rosier, M., Vandekerckhove, J., and Rosseneu, M. (2003) *J. Mol. Biol.* **325,** 259–274
2. Schmitz, G., and Kaminski, W. E. (2002) *Curr. Atheroscler. Rep.* **4,** 243–251
3. Brooks-Wilson, A., Marcil, M., Clee, S.M., Zhang, L.H., Roomp, K., van Dam, M., Yu, L., Brewer, C., Collins, J.A., Molhuizen, H. O., Loubser, O., Ouelette, B.F., Fichter, K., Ashbourne-Excoffon, K. J., Sensen, C. W., Scherer, S., Mott, S., Denis, M., Martindale, D., Frohlich, J., Morgan, K., Koop, B., Pimstone, S., Kastelein, J. J., and Hayden, M. R. (1999) *Nat. Genet.* **22,** 336–345
4. Bodzioch, M., Orso, E., Klucken, J., Langmann, T., Bottcher, A., Diederich, W., Drobnik, W., Barlage, S., Buchler, C., Porsch-Ozcurumez, M., Kaminski, W. E., Hahmann, H. W., Oette, K., Rothe, G., Aslanidis, C., Lackner, K. J., and Schmitz, G. (1999) *Nat. Genet.* **22,** 347–351
5. Rust, S., Rosier, M., Funke, H., Real, J., Amoura, Z., Piette, J. C., Deleuze, J. F., Brewer, H. B., Duverger, N., Denefle, P., and Assmann, G. (1999) *Nat. Genet.* **22,** 352–355
6. Chen, W., Sun, Y., Welch, C., Gorelik, A., Leventhal, A. R., Tabas, I., and Tall, A. R. (2001) *J. Biol. Chem.* **276,** 43564–43569
7. Higgins, C. F. (1992) *Annu. Rev. Cell Biol.* **8,** 67–113
8. Holland, I. B., and Blight, M. A. (1999) *J. Mol. Biol.* **293,** 381–399
9. Dean, M., Hamon, Y., and Chimini, G. (2001) *J. Lipid Res.* **42,** 1007–1017
10. Allikmets, R. (2000) *Am. J. Hum. Genet.* **67,** 793–799
11. Schmitz, G., and Langmann, T. (2001) *Curr. Opin. Lipidol.* **12,** 129–140
12. Winter, M. C., and Welsh, M. J. (1997) *Nature* **389,** 294–296
13. Picciotto, M. R., Cohn, J. A., Bertuzzi, G., Greengard, P., and Nairn, A. C. (1992) *J. Biol. Chem.* **267,** 12742–12752
14. Vanoye, C. G., Castro, A. F., Pourcher, T., Reuss, L., Altenberg, G. A. (1999) *Am. J. Physiol.* **276,** C370–C378
15. See, R. H., Caday-Malcolm, R. A., Singaraja, R. R., Zhou, S., Silverston, A., Huber, M. T., Moran, J., James, E. R., Janoo, R., Savill, J. M., Rigot, V., Zhang, L.H., Wang, M., Chimini, G., Wellington, C. L., Tafuri, S. R., and Hayden, M. R. (2002) *J. Biol. Chem.* **277,** 41835–41842
16. Yamauchi, Y., Hayashi, M., Abe-Dohmae, S., and Yokoyama, S. (2003) *J. Biol. Chem.* **278,** 47890–47897
17. Martinez, L. O., Agerholm-Larsen, B., Wang, N., Chen, W., and Tall, A. R. (2003) *J. Biol. Chem.* **278,** 37368–37374
18. Sarno, S., Moro, S., Meggio, F., Zagotto, G., Dal Ben, D., Ghisellini, P., Battistutta, R., Zanotti, G., and Pinna, L. A. (2002) *Pharmacol. Ther.* **93,** 159–168
19. Litchfield, D. W. (2003) *Biochem. J.* **369,** 1–15
20. Meggio, F., and Pinna, L. A. (2003) *FASEB J.* **17,** 349–368
21. Roosbeek, S., Vanloo, B, Duverger, N., Caster, H., Breyne, J., De Beun, I., Patel, H., Vandekerckhove, J., Shoulders, C., Rosseneu, M., and Peelman, F. (2001) *J. Lipid Res.* **42,** 31–40
22. Lopez, J., Latta, M., Collet, X., Vanloo, B., Jung, G., Denefle, P., Rosseneu, M., and Chambaz, J. (1994) *Eur. J. Biochem.* **225,** 1141–1150
23. Rigot, V., Hamon, Y., Chambenoit, O., Alibert, M., Duverger, N., and Chimini, G. (2002) *J. Lipid Res.* **43,** 2077–2086
24. Roosbeek, S., Caster, H., Zhou, Liu Q., Berne, P. F., Duverger, N., Christiaens, B., Vandekerckhove, J., Peelman, F., Labeur, C., and Rosseneu, M. (2004) *Protein Expression Purif.* **35,** 102–110
25. Filhol, O., Cochet, C., Wedegaertner, P., Gill, G. N., and Chambaz, E. M. (1991) *Biochemistry* **30,** 11133–11140
26. Amoresano, A., Cirulli, C., Quemeneur, E., and Marino, G. (2004) *Eur. J. Mass Spectrom.* **10,** 401–412
27. Chambenoit, O., Hamon, Y., Marguet, D., Rigneault, H., Rosseneu, M., and Chimini, G. (2001) *J. Biol. Chem.* **276,** 9955–9960
28. Cheng, H. H., Xu, L., Kumagai, H., and Simoni R. D. (1999) *J. Biol. Chem.* **274,**

17171–17178
29. Harkin, D. P., Bean, J. M., Miklos, D., Song, Y. H., Truong, V. B., Englert, C., Christians, F. C., Ellisen, L. W., Maheswaran, S., Oliner, J. D., and Haber, D. A. (1999) *Cell* **97,** 575–586
30. Oram, J. F., Lawn, R. M., Garvin, M. R., and Wade, D. P. (2000) *J. Biol. Chem.* **275,** 34508–34511
31. Miller, S. J., Lou, D. Y., Seldin, D. C., Lane, W. S., and Neel, B. G. (2002) *FEBS Lett.* **528,** 145–153
32. Chang, G., and Roth, C. B. (2001) *Science* **293,** 1793–1800
33. Locher, K. P., Lee, A. T., and Rees, D. C. (2002) *Science* **296,** 1091–1098
34. Van Veen, H. W., Higgins, C. F., and Konings, W. N. (2001) *J. Mol. Microbiol. Biotechnol.* **3,** 185–192
35. Martel, V., Filhol, O., Nueda, A., and Cochet, C. (2002) *Ann. N. Y. Acad. Sci.* **973,** 272–277
36. Maizel, A., Tassetto, M., Filhol, O., Cochet, C., Prochiantz, A., and Joliot, A. (2002) *Development* **129,** 3545–3553
37. Goodfellow, H. R., Sardini, A., Ruetz, S., Callaghan, R., Gros, P., McNaughton, P. A., and Higgins, C. F. (1996) *J. Biol. Chem.* **271,** 13668–13674

Mapping phosphorylation sites: a new strategy based on the use of isotopically-labelled dithiothreitol and mass spectrometry

Angela Amoresano and Gennaro Marino*
Department of Organic Chemistry and Biochemistry, Federico II University of Naples, Naples, Italy and School of Biotechnological Sciences, Federico II University of Naples, Naples, Italy

Claudia Cirulli
Department of Organic Chemistry and Biochemistry, Federico II University of Naples, Naples, Italy

Eric Quemeneur
Department of Protein Engineering, Life Science Division, C.E.A., Marcoule, France

Phosphoproteomics, nowadays, represents a front line in functional proteomics as testified by the number of papers recently appearing in the literature. In an attempt to improve and simplify the methods so far suggested we have set up a simple isotope-coded approach to label and quantitate phospho-Ser/-Thr residues in protein mixtures. First of all, after appropriate oxidation of cysteine/cystine residues followed by tryptic hydrolysis, we have optimised and simplified the β-elimination reaction to get the corresponding alkene moiety from the phosphate esters. This was achieved by (a) separating the elimination reaction from the addition reaction, (b) the use of $Ba(OH)_2$ as alkali reagent and (c) its further elimination by the simple addition of solid CO_2 to the peptide mixture. The Michael reaction was then performed, after the removal of $BaCO_3$ by centrifugation, by adding dithiothreitol (DTT) to the peptide mixture. Finally, the direct purification of the modified phosphopeptides was performed on a thiol-sepharose column. The availability of fully deuterated DTT, introducing a 6 Da difference with respect to the non-deuterated species, allows quantitation of the differential extent of signalling modification when analysed by matrix-assisted laser desorption/ionisation mass spectrometry (MALDI-MS) and liquid chromatography/mass spectrometry. The entire procedure has been set up by using bovine α-casein, and resulted in the identification of all the phosphorylated tryptic peptides, including the tetraphosphorylated peptides, which escaped all previously reported procedures.

Keywords: α-casein, phosphoproteome, ICAT, mass spectrometry

Introduction

In the living cell, many processes are governed not only by the relative abundance of proteins but also by rapid and transient regulation of the activity, association and localization of proteins and protein complexes. Dynamic post-translational modification is a general mechanism for maintaining and regulating protein structure and function. One of the most important protein modifications that highly influences protein activity and propagates signals within cellular pathways is phosphorylation.[1] Protein phosphorylation is recognized as the most important action in the regulation of cell functions. Although regulation can occur at the level of protein synthesis along with concomitant changes in the extent of phosphorylation, there are also examples of regulation of protein function by phosphorylation without altering protein abundance. Therefore, the precise characterization of the subset of proteins in the proteome that becomes modified *in vivo* by phosphorylation (phosphoproteome) as well

**Corresponding author: Professor Gennaro Marino, Dipartimento di Chimica Organica e Biochimica, Complesso Universitario Monte Sant' Angelo, Via Cinthia 6, I-80126 Napoli, Italy. E-mail: gmarino@unina.it.*

DOI: 10.1255/ejms.599 ISSN 1356-1049

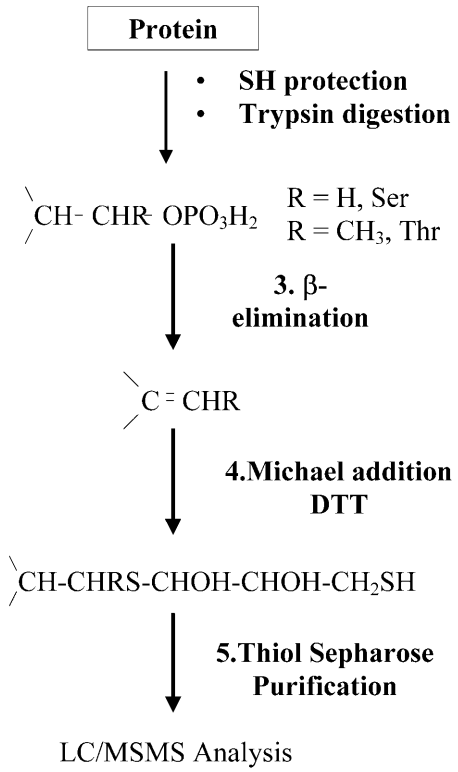
as the localization and quantitation of the phosphorylation sites are essential to fully understand how proteins are activated or inhibited within the signal-transduction pathway, thus allowing speculations to be made about the set of kinases and phosphatases involved.[2–6]

Unfortunately, information on the protein phosphorylation cannot be deduced from gene sequences. Techniques to identify the phosphorylation site by ^{32}P-labelling coupled with Edman sequence analysis are generally thought to be difficult and not always precise. Phosphoamino acid antibodies have been exploited recently but their use has actually been restricted to tailor-made immunological applications.[7–10] Another approach involves the use of immobilized metal–ion affinity chromatography (IMAC) to selectively retain and enrich phosphopeptides from mixtures.[11-13] However, non-phosphorylated peptides also can be isolated through non-specific interaction thus complicating subsequent analyses.

Mass spectrometry is widely employed nowadays as a key tool in the analysis of protein phosphorylation[14–21] since the pioneering work of Pucci and co-workers.[22] Recent advances in proteomics procedures and technologies have led to high-throughput protein identification in a complex protein mixture. New methods that involve modifying phosphoproteins with affinity tags, in combination with stable isotope incorporation,[23–27] have been developed for the specific enrichment and quantitation of phosphopeptides. Gygi and co-workers have recently reported[28] an approach for quantitative differential analysis that employs isotope-coded affinity tags (ICAT). This method involves the affinity-isolation of cysteine-containing peptides by modifying proteins with a Cys-specific reagent that contains a linker arm connecting a thiol group to a biotin moiety. The ICAT reagent is used in both "light" and "heavy" forms: eight hydrogen atoms in the linker arm have been substituted by eight deuterium atoms in the heavy version of the reagent. Derivatization of two distinct proteomes with the two versions of the ICAT reagent provides a basis for protein quantitation.

In this paper, we describe an improved approach, based on simple chemical manipulations and mass spectrometric procedures (see Scheme 1), for the selective analysis of phosphoserine and phosphothreonine in protein mixtures following conversion of the peptide phosphate moiety into dithiothreitol (DTT) derivatives. The new residues in the sequence were shown to be stable and easily identifiable under the general conditions for tandem mass spectrometry sequencing, applicable to the precise localization of the exact site of phosphorylation. The suitability of the method with stable-isotope coding is also demonstrated.

While this work was in progress, a paper by Wells[29] described an almost identical approach for selective labeling of O-glycosylation sites. However, the major aim of the present work is devoted to the use of isotopically-labelled DTT, thus allowing a simple and direct quantitative analysis by mass spectrometry.



Scheme 1.

Results and discussion

This paper reports a novel approach to phosphoprotein mapping, based on site-specific modification of phosphoseryl/phosphothreonyl residues. The validity of its applicability for the identification of the phosphorylation sites in peptides was tested using a matrix-assisted laser/desorption ionization (MALDI) tandem time-of-flight mass spectrometer (ToF/ToF-MS). Here, we have shown that an alternative conversion of the P-Ser and P-Thr residues into compounds, which are stable during collision-induced dissociation (CID), provides easily interpretable product-ion spectra.

To test our procedure under realistic but controlled conditions, we carried out the reactions by using phosphorylated bovine α-casein, a normal protein utilized in these studies.[24,25,28] The protein samples after separation on sodium dodecyl phosphate (SDS) polyacrylamide gel were reduced, alkylated and digested *in situ* with trypsin (see Materials and Methods section). The peptide mixtures were then directly analysed by MALDI-MS (Table 1). The phosphopeptides were identified by their 80 Da mass difference compared with the peptides expected from the non-phosphorylated sequence. As an example, the signal at *m/z* 1951.9 in Figure 1 was assigned to the peptide 104–119 within the α-casein sequence, but carrying a phosphate group. The peptide mixture was submitted to the strategy described above (see Scheme 1) and each step was monitored via MALDI-MS. Reaction conditions were optimized for both the β-elimination and DTT-addition reactions.

Table 1. MALDI-MS analysis of the peptide mixture of bovine alpha-casein digested with trypsin.

Experimental MH^+	Theoretical MW	Peptide	Note
3026.3	3025.4	166–193	
2316.31	2315.13	133–151	
1952.3	1950.94	104–119	1P
1759.7	1758.93	8–22	
1660.5	1659.8	106–119	1P
1384.5	1383.73	23–34	
1337.7	1336.6	80–90	
1267.5	1266.69	91–100	
946.3	945.5	35–42	
910.5	909.46	125–132	
875.6	874.55	1–7	
831.3	830.37	84–90	
748.5	747.3	194–199	
615.38	614.32	120–124	

Note reports the number of phosphate groups (P) added to the peptides

β-*Elimination reaction*

Under strongly alkaline conditions, the phosphate moiety on P-Ser and P-Thr undergoes β-elimination to form dehydroalanine (ΔSer) or dehydro-2-aminobutyric acid (ΔThr), respectively. In the case of O-glycosylated peptides, the corresponding glycopeptides have to be separated by a simple lectin-affinity chromatography step. The α,β-unsatured residues are potent Michael acceptors, which can readily react with a nucleophile.[30,31]

In the present study, the phosphate moieties were removed via barium hydroxide ion-mediated β-elimination from P-Ser and P-Thr. The barium hydroxide was used in place of the previously reported NaOH[32,33] taking in account the higher purity of this reagent and its higher reactivity towards the P-Thr residues. To facilitate removal of excess reagent, we simply made use of solid carbon dioxide. The use of this reagent permitted the removal of barium carbonate as a pellet on the bottom of a vial after centrifugation. It is worth noting that when the pellet of barium carbonate was dissolved and analysed by MALDI-MS, no peptide was detected in the mass spectrum, thus indicating the absence of any sort of non-specific precipitation. On the contrary, the mass spectral analysis carried out on the supernatant showed the occurrence of a series of signals corresponding to the β-eliminated peptides. As an example, the signal at *m/z* 1854.3 was assigned to the peptide 104–119 of α-casein since it occurs 98 Da lower than the mass recorded for the phosphorylated peptide, thus indicating that P-Ser115 was converted to dehydroalanyl by the β-elimination reaction (Figure 2).

DTT as a bifunctional reagent

The β-eliminated peptide mixture was submitted to a Michael-type addition by using dithiothreitol (DTT) and following the procedure described.[30] The rationale behind the choice of this reagent is due to its higher solubility and reactivity compared with the already proposed 1,2-ethanedithiol (EDT).[31–33] In fact, EDT has limited solubility in aqueous medium, thereby requiring sample clean-up prior to liquid chromatography (LC)/MS analysis since the conversion is carried out in the presence of a substantial amount of organic solvent.[34] The addition reaction results in the creation of a free thiol group in place of what was formerly a phosphate moiety. The extent of reaction was monitored using MALDI-MS and indicated that the yield from the addition reaction of a DTT-modified peptide was about 70% (Figure 3), as expected for a typical Michael reaction addition.[30] A comparison made using EDT showed that, even in the presence of a very large reagent excess, yields were not higher than 30% (data not shown). However, the direct mass spectral analysis of a complete peptide mixture suffered from suppression phenomena leading to the non-appearence of signals due to most of the phosphorylated peptides.

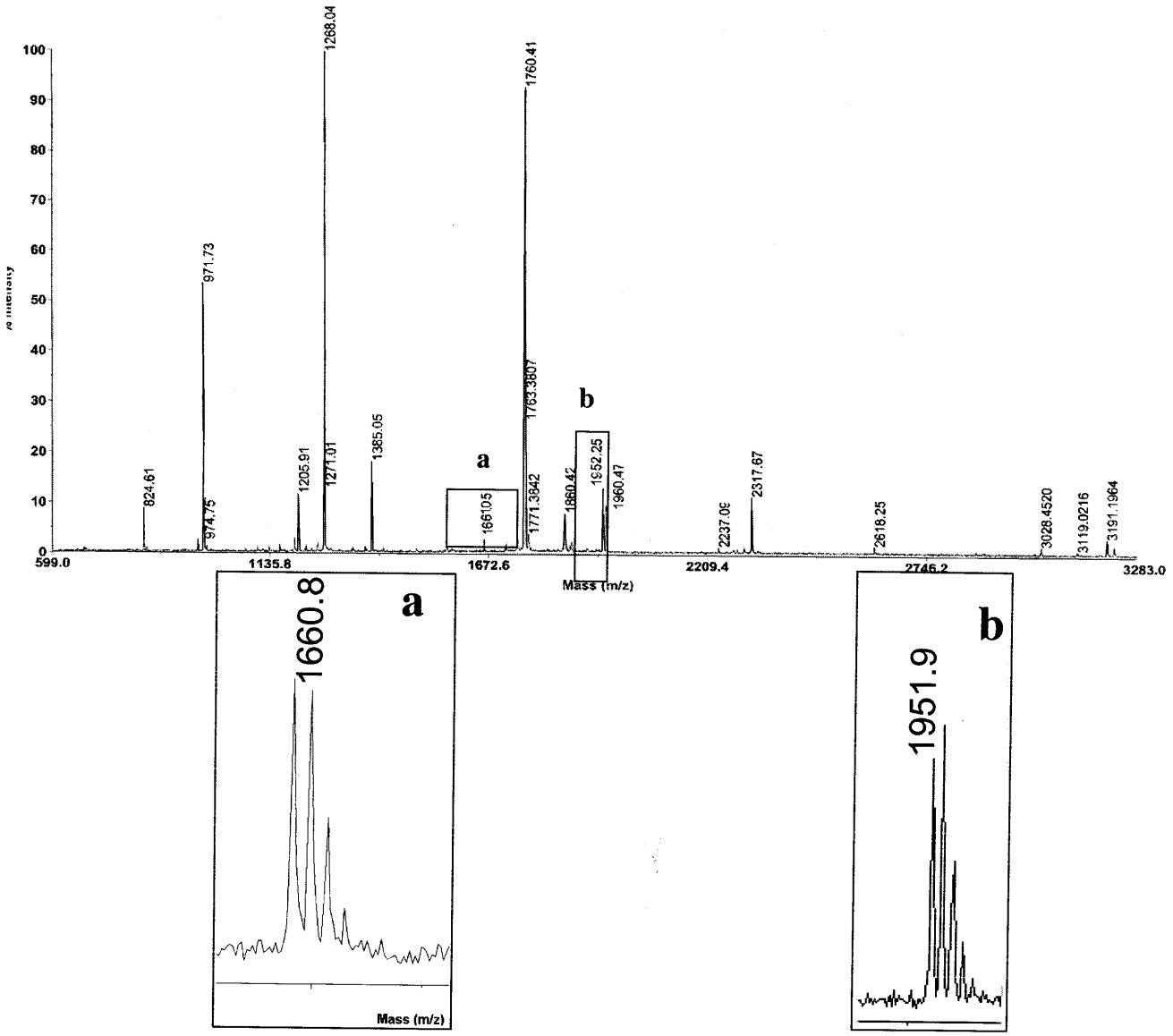


Figure 1. MALDI-MS analysis of the trypsin peptide mixture from α-casein. The insets report the amplification of the region of the spectrum showing the phosphorylated peptides.

Affinity capture of thiolated peptides

Specific enrichment of thiolated peptides is easily achieved using a Sepharose thiopyridyl resin. MALDI-MS analysis performed on the peptides, eluted from the affinity purification step, now reveals major ion signals at *m/z* 1716.9, 2008.3, 2039.1 and 3001.2. These values were assigned to the β-eliminated fragments 106–119, 104–119, 43–58 and 59–79 modified by 1, 1, 2 and 5 DTT adducts, respectively, as indicated in Figure 4. It is worth noting that the described procedure resulted in the identification of all the phosphorylated tryptic peptides, including the pentaphosphorylated peptide that has so far escaped detection in different attempts.[24,25,28] The data are summarized in Table 2.

Tandem mass spectrometry (MS/MS) analysis

Sequence information can be obtained either by MALDI-ToF in the post-source decay (PSD) mode or by MALDI-ToF/ToF. The precise localization of the phosphorylation sites in the peptides was achieved by MS/MS sequencing. The conversion of P-Ser and P-Thr into DTT-tagged residues provides compounds that are stable under high-energy collision in the MALDI-ToF/ToF instrument. In fact, the tandem mass spectral analyses by MALDI-ToF/ToF performed on the precursor ion at *m/z* 2008.5 (Figure 5) from the tryptic mixture of α-casein after labelling with DTT revealed the presence of diagnostic fragment ions. The derivatized product ions were identified from either the corresponding b- or y-fragment ion series. The MS/MS

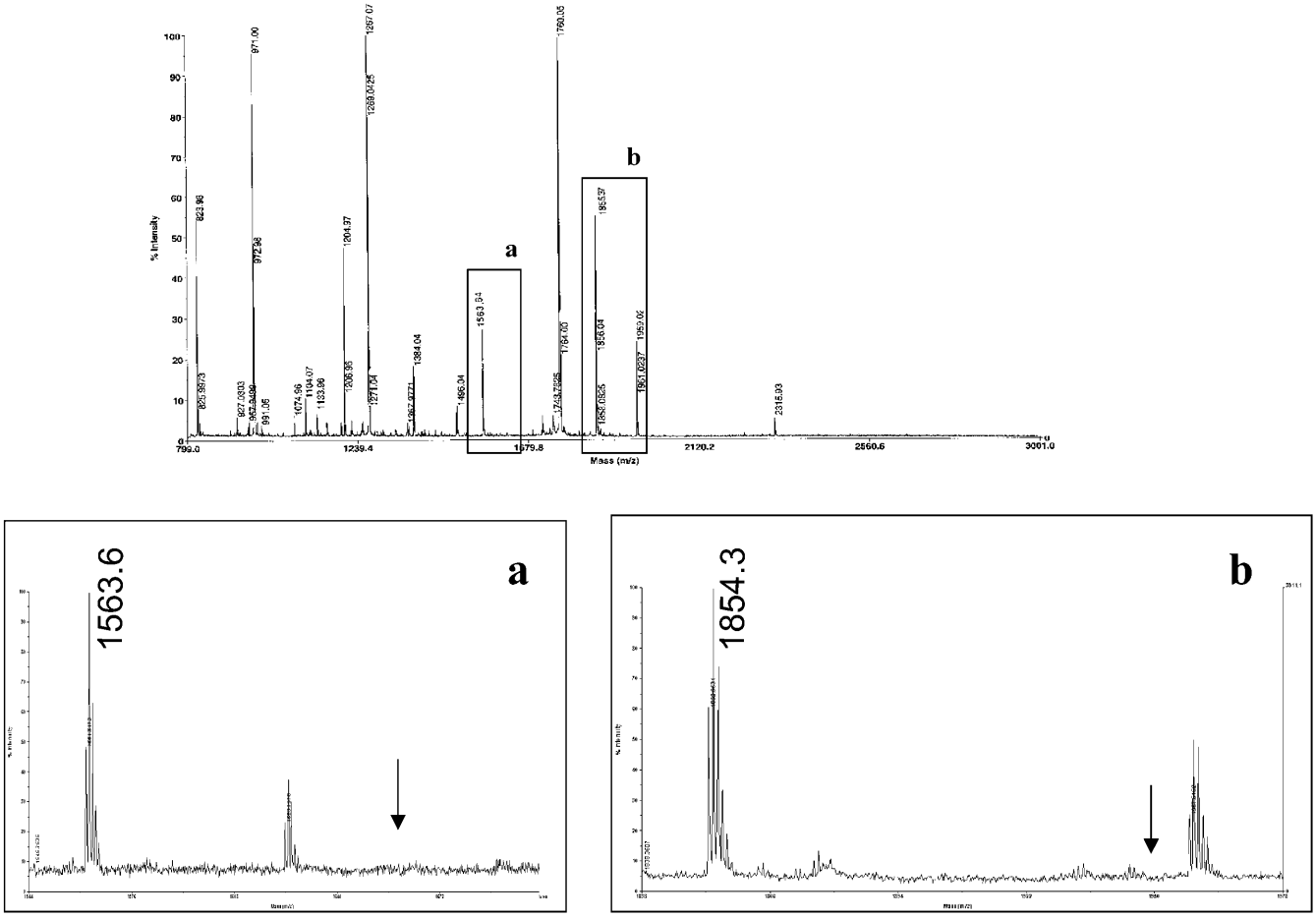


Figure 2. MALDI-MS analysis of trypsin peptide mixture from α-casein following the β-elimination reaction. The boxes report amplifications of the region of the spectrum showing the β-eliminated peptides. The arrows indicate the disappearance of the phosphorylated peptides and the occurrence of the new β-eliminated signals occurring at minus 98 Da.

spectrum revealed the complete series of N-terminal fragment ions (b1–b15) plus the C-terminal ions (y1–y16) for the modified phosphopeptide with a DTT-labelled phosphoserine residue. The fragment ions y4 and y5, in fact, show a mass difference of 223 Da corresponding to a β-eliminated serine residue, labelled with DTT, thus confirming the phosphorylation site as Ser115.

Differential isotope-coded analysis

To address the problem of quantification,[28,35,36] having created a thiol group at the phosphoserine/phosphothreonine site, it is possible to follow, basically, the strategy outlined by Gygi *et al.*[28] However, because of the availability of fully-deuterated DTT, we exploited a much simpler isotope-tagging strategy. The peptides were quantified by measuring, in the mixture, the relative signal intensities for pairs of peptide ions of identical sequence differentially labelled with light or heavy DTT. The mass difference between the light and heavy DTT reagent molecules (6 Da) reliably provides a good separation between the molecular ions. As an example, the mass signals recorded in the MALDI spectrum, at *m/z* 2008.5 and 2014.5, were attributed to the DTT-labelled peptide 104–119, modified with the light DTT and heavy DTT forms, respectively. Pairs of peptides tagged with the light DTT and heavy DTT reagents, respectively, are chemically identical and, therefore, serve as ideal mutual internal standards for quantitation. The ratio between the intensities of the lower and upper mass components of these peaks provides an accurate measurement of the relative abundance of each peptide and, hence, of the related proteins in the original cell pool because the relative intensity in the mass spectrum of a given peptide is independent of the isotopic composition of the defined, isotopically-tagged reagent.

To further illustrate the ability of this labelling strategy in quantifying the relative phosphorylation of a peptide from two different samples, DTT and DTT-D6 were used to label samples containing stoichiometric concentrations of α-casein in ratios of 1 : 1, 2 : 1 and 4 : 1. The peptide mixtures were analysed by MALDI-MS operating in reflectron mode. The spectra were obtained by acquiring about 2000 shots per

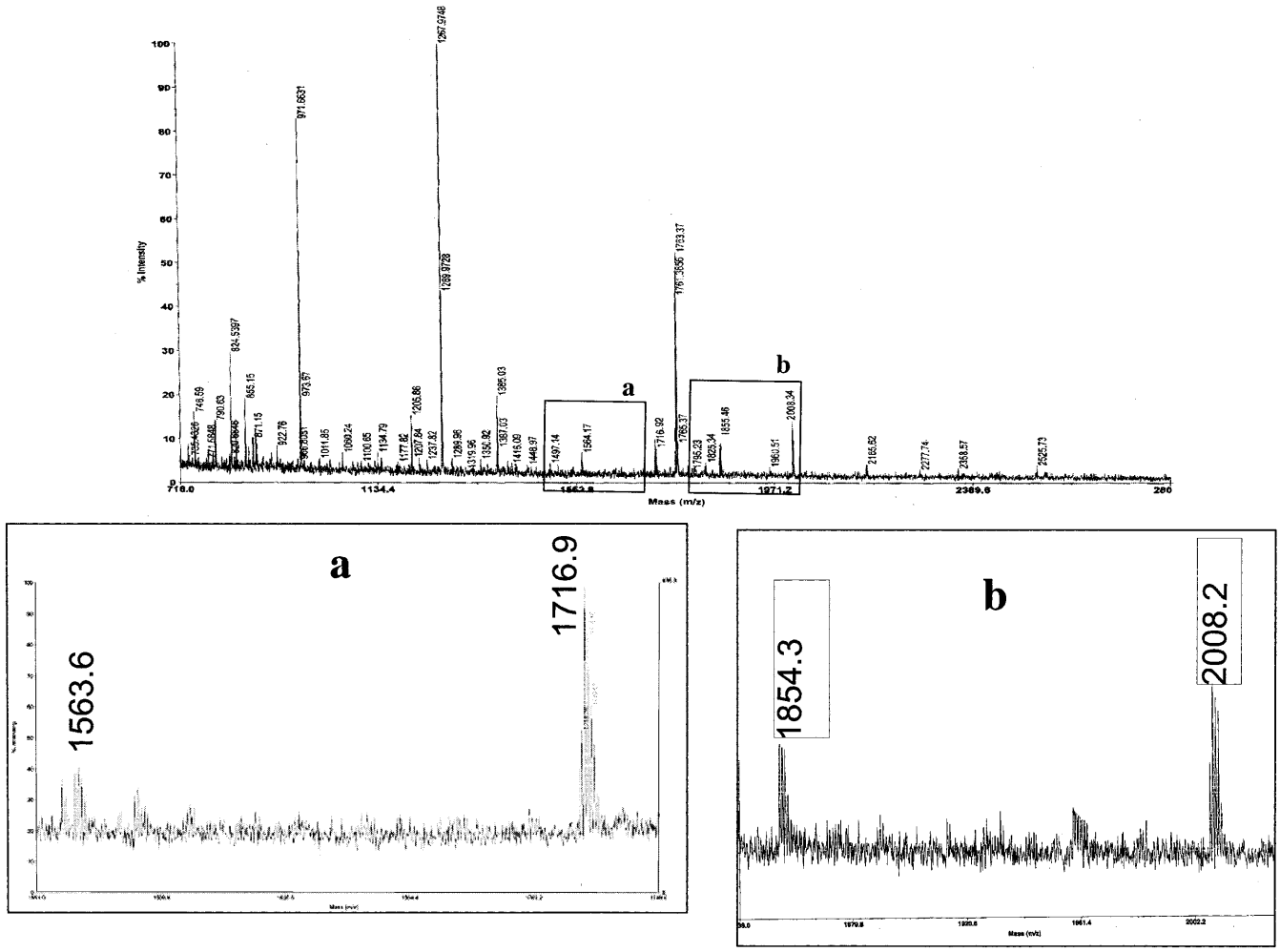


Figure 3. MALDI-MS analysis of the trypsin peptide mixture from the protein after DTT labelling reaction. Insets: enlarged region of the spectrum showing the β-eliminated and DTT-labelled peptides.

spectrum in order to produce a "stable" peak. In our experience, this is the typical number of shots giving an even distribution of shots on the entire well surface, thus allowing a confident averaging independently of the peptide content of the single crystals. As shown for the DTT peptide 104–109 (Figure 6), the ratios were automatically integrated for each mass spectrum, resulting in excellent agreement with the concentration used in the labelling experiment (Table 3). An aliquot of the same peptide mixture was submitted to a more rigorous quantitative analysis via LC/MS (Figure 7), as reported in Table 4, demnnstrating the correctness of the previous MALDI-MS data.

Table 2. MALDI-MS analysis of the purified DTT-labelled peptides of the tryptic digest of α-casein.

Experimental MH^+	Theoretical MW	Peptide	Note
3001.2	2999.9	59–79	5 DTT
2039.1	2038.68	43–58	2 DTT
2008.3	2006.94	104–119	1 DTT
1716.9	1715.8	106–119	1 DTT

Note reports the number of dithiothreitol (DTT) groups added to the peptides

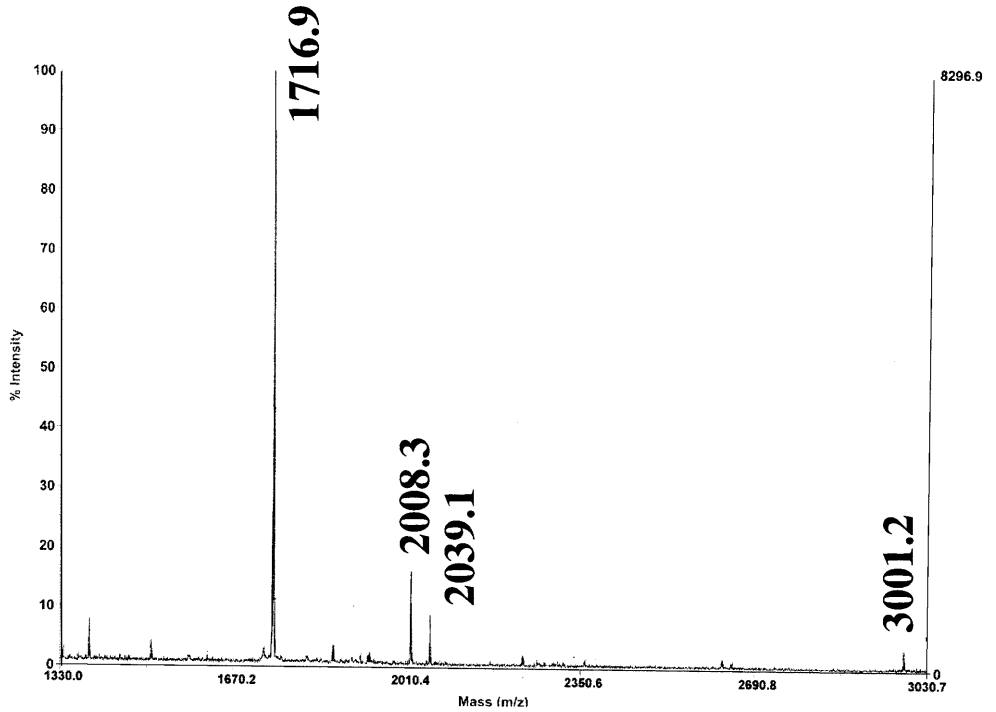


Figure 4. MALDI-MS analysis of DTT-labelled peptides following purification step. Each signal is attributed to the corresponding modified peptide according to Table 2.

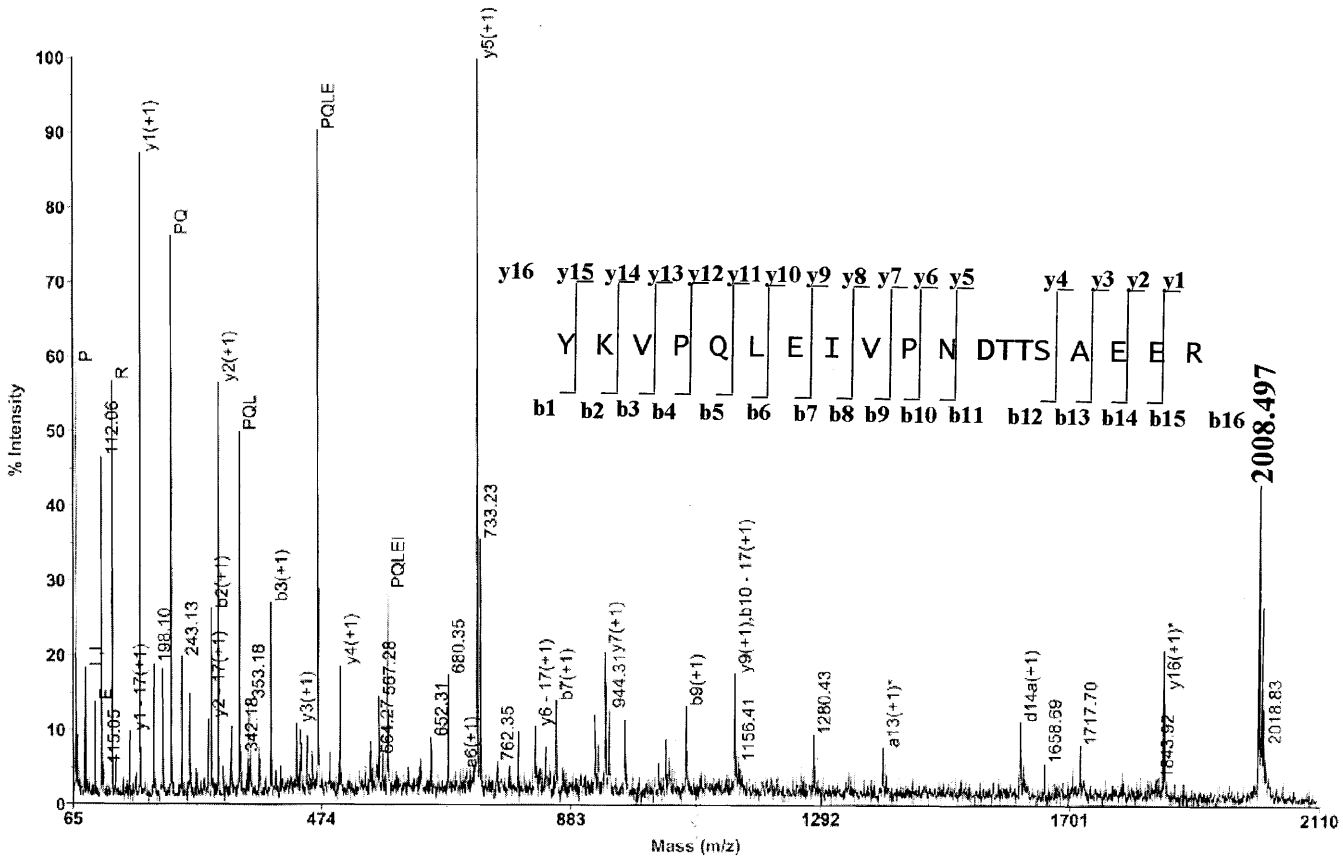


Figure 5. Tandem mass spectrometry identification of α-casein phosphopeptide. The Ser residue containing the DTT label is identified by the loss of *m/z* 223.4 between the y4- and y5-ions of the modified peptide.

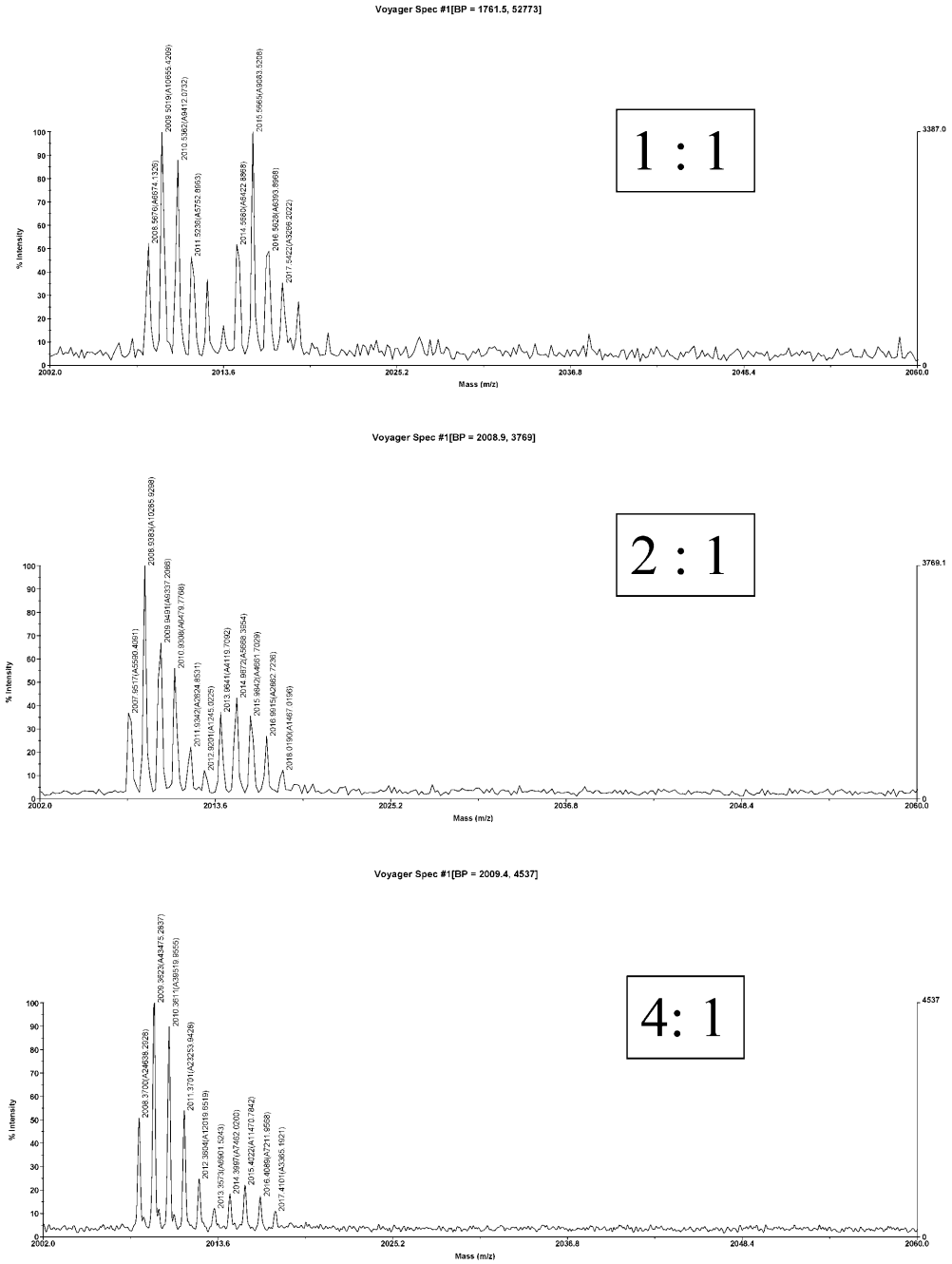


Figure 6. Stoichiometric conversion of the phosphorylated states using DTT-D0/D6 labelling via MALDI-MS. Partial MALDI-MS spectra of the modified peptide 104–119. Samples of α-casein containing ratios of 1 : 1, 2 : 1 and 4 : 1 were labelled with DTT-D0/DTT-D6, combined, and isolated using the DTT strategy reported in Scheme 1 and directly analysed via MALDI-MS.

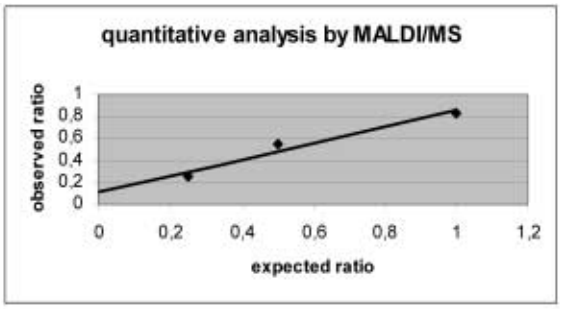
Conclusions

This paper reports a simple integrated methodology, which involves chemical replacement of the phosphate moitties by affinity tags. This method for locating and quantitating phosphorylated residues should provide results of interest in signalling pathways and control mechanism studies involving phosphorylation or dephosphorylation of serine/threonine residues. Work is presently in progress on samples of different origin (bacterial, yeast, human) to validate this simple methodology set up on real examples.

It is worth considering that such a quantitative approach can be extended to the labelling of O-glycosylation sites. In fact, the O-glycopeptides retained on a lectin-affinity column can be eluted and submitted to the procedure described by Wells and co-workers.[29] The quantitative analysis can be performed by using light and heavy forms of DTT to selectively label the modified sites, thus leading to a

Table 3. Quantitative analysis by MALDI-MS.

Expected ratio	DTT (D0-labelled) peak area	DTT (D6-labelled) peak area	Observed ratio
1 : 1	29168.1	24306.7	1.2 : 1
2 : 1	34318.05	18779.51	1.8 : 1
4 : 1	142907.11	36411.46	3.9 : 1



$Y=0.7429x + 0.11$ $R^2 = 0.9565$

quantitative description of O-glycosylation related to different cell growth conditions.

Materials and methods

Materials and reagents

TPCK-treated trypsin, dithiothreitol, barium hydroxide and α-cyano-4-hydroxycinnamic acid were from Sigma–Aldrich (St Louis, MO, USA). Deuterated dithiothreitol was purchased from CEA (Saclay, France). Sepharose thiopyridyl resin was from Pierce Biotechnology (Rockford, IL, USA). All other reagents and solvents were of the highest purity available from Carlo Erba (Milan, Italy).

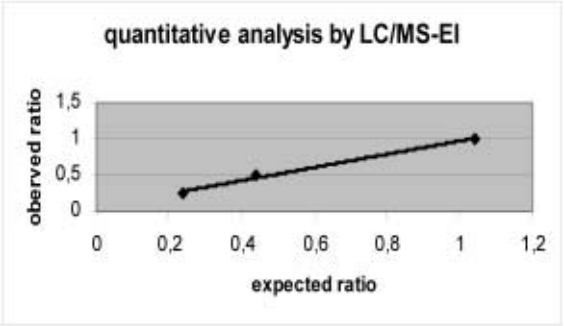
In-situ *digestion*

Proteins were fractionated by SDS-electrophoresis on 12.5% polyacrylamide gels under reducing conditions.

The analysis was performed on the Comassie blue-stained proteins excised from gels. The excised spots were

Table 4. Quantitative analysis by LC/ESI-MS

Expected ratio	DTT (D0) labelled peak area	DTT (D6) labelled peak area	Observed ratio
1 : 1	382829	396550	0.96 : 1
2 : 1	1312411	578038	2.27 : 1
4 : 1	2134041	505166	4.22 : 1



$Y = 0.9135x + 0.0596$ $R^2 = 0.9918$

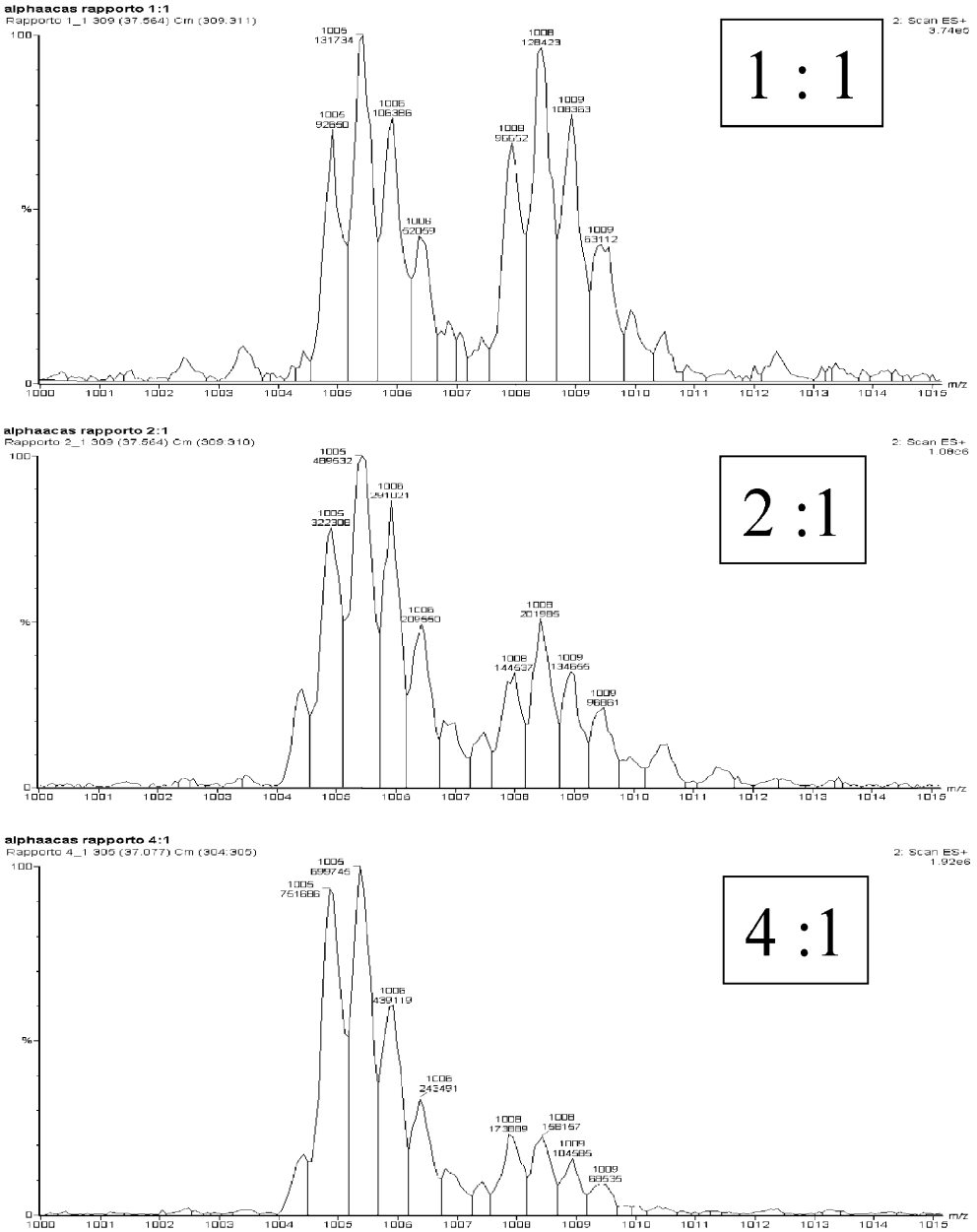


Figure 7. Stoichiometric conversion of the phosphorylated states using DTT-D0/D6 labelling and electrospray (ES)MS. The LC/ESMS spectrum shows the doubly-charged ion of the modified peptide 104–119. Samples of α-casein containing ratios of 1 : 1, 2 : 1 and 4 : 1 were labelled with DTT-D0:DTT-D6, combined, and isolated using the DTT strategy reported Scheme 1 and analysed via LC/ESMS.

washed first with acetonitrile (ACN) and then with 0.1 M ammonium bicarbonate. Protein samples were reduced by incubation in 10 mM dithiothreitol (DTT) for 45 min at 56°C. The cysteines were alkylated by incubation in 5 mM iodoacetamide for 15 min at room temperature in the dark. The gel particles were then washed with ammonium bicarbonate and ACN.

Enzymatic digestion was carried out with trypsin (12.5 ng μL^{-1}) in 50 mM ammonium bicarbonate pH 8.5 at 4°C for 4 h. The buffer solution was then removed and a new aliquot of the enzyme/buffer solution was added for 18 h at 37°C. A minimum reaction volume, enough for the complete rehydratation of the gel was used. Peptides were extracted by washing the gel particles with 20 mM ammonium bicarbonate and 0.1% trifluoroacetic acid (TFA) in 50% ACN at room temperature and then lyophilized.

Phosphate group modification

β-Elimination reactions were carried out by incubating the peptide mixture in 55 M $Ba(OH)_2$, at 37°C, for 90 min under nitrogen. Solid carbon dioxide was then added. The precipitated barium carbonate was then removed by

centrifugation at 13,000 g for 5 min. Then DTT, in the light and heavy (fully-deuterated) form, 30% w/v in 10 mM Hepes buffer (pH 7.5), was added during 3 h, at 50°C under nitrogen.

Isolation and enrichment of tagged peptides

After the addition reaction, the peptide mixtures were treated with activated thiol sepharose resin directly in batch. The resin was first washed in water (× 2) and then with 0.1 M Tris-HCl binding buffer, (pH 7.5). The peptide mixture was added to the resin and washed with the binding buffer. The modified peptides were then extracted using 20 mM DTT in 10 mM Tris-HCl elution buffer (pH 7.5).

Mass spectrometric analyses

MALDI mass spectra were recorded using an Applied Biosystem Voyager DE-PRO instrument operated in the reflectron mode. A mixture of analyte solution and α-cyano-hydroxycinnamic acid [10 mg mL^{-1} in acetonitrile + ethyl alcohol + 0.1% trifluoroacetic acid (1 : 1 : 1, v / v/ v)] was applied to the metallic sample plate and dried under vacuum. Mass calibration was performed with insulin at *m/z* 5734.5 and a matrix peak at *m/z* 379.3 as internal standards. Raw data were analysed by using computer software provided by the manufacturers and reported as average masses. MS/MS spectra were recorded using an Applied Biosystem 4700 Proteomics Analyzer mass spectrometer operating in reflectron mode.

Peptide mixtures were analysed by LC-MS on a single-quadrupole ZQ electrospray instrument (Waters Micromass) coupled to an HPLC (2690 Alliance purchased by Waters) on a Phenomenex 30 mm × 0.46 mm i.d. reversed-phase column. Peptides from the mixture were eluted at a flow rate of 1 mL min^{-1} with a 10–95% acetonitrile-water gradient over 60 min. Data were acquired and processed by centroiding an isotopic distribution using the Mass-Lynx program (Micromass).

References

1. R.G. Krishna and F. Wold, "Posttranslational modification of proteins", *Adv. Enzymol. Mol. Biol*. **67,** 265 (1993).
2. T. Hunter, "Signalling–2000 and beyond", *Cell* **100,** 113 (2000).
3. X. Feng and Y.A. Hannun, "An essential role for autophosphorylation in the dissociation of activated protein kinase C from the plasma membrane", *J. Biol. Chem.* **273,** 26870 (1998).
4. S.H. Francis and J.D. Corbin, "Cyclic AMP and cyclic GMP in cell signalling", in *Signal Transduction,* Ed by C.H. Heldin and M. Purton. Chapman and Hall, London, UK, p. 223 (1996).
5. D.B. Parekh, W. Ziegler and P.J. Parker, "Multiple pathways control protein kinase C phosphorylation", *EMBO J.* **19,** 496 (2000).
6. D.F. Stojdl and J.C. Bell, "Protein kinases: the splice of life", *Biochem. Cell. Biol.* **77,** 293 (1999).
7. A. Sickmann and H.E. Meyer, "Phosphoamino acid analysis", *Proteomics* **1,** 200 (2001).
8. S. Sullivan and T.W. Wong, "A manual sequencing method for identification of phosphorylated amino acids in phosphopeptides", *Anal. Biochem.* **197,** 65 (1991).
9. J.X. Yan, N.H. Packer, A.A. Gooley and K.L. Williams, "Protein phosphorylation: technologies for the identification of phosphoamino acids", *J. Chromatogr. A* **808,** 23 (1998).
10. X.L. Zhang, C.J. Herring, P.R. Romano, J. Szczepanowska, H. Brzeska, A.G. Innebusch and J. Qin, "Identification of phosphorylation sites in proteins separated by polyacrylamide gel electrophoresis", *Anal. Chem.* **70,** 2050 (1998).
11. L. Andersson and J. Porath, "Isolation of phosphoproteins by immobilized metal (Fe3+) affinity chromatography", *Anal. Biochem.* **154,** 250 (1986).
12. P. Cao and J.T. Stults, "Phosphopeptide analysis by on-line immobilized metal-ion affinity chromatography-capillary electrophoresis-electrospray ionization mass spectrometry", *J. Chromatogr. A* **853,** 225 (1999).
13. M.C. Posewitz and P. Tempst, "Immobilized gallium (III) affinity chromatography of phosphopeptides", *Anal. Chem.* **71,** 2883 (1999).
14. J.M. Asara and J. Allison. "Enhanced detection of phosphopeptides in matrix-assisted laser desorption/ionization mass spectrometry using ammonium salts", *J. Am. Soc. Mass Spectrom.* **10,** 35 (1999).
15. M.E. Belov, M.V. Gorshkov, H.R. Udseth, G.A. Anderson and R.D. Smith, "Zeptomole-sensitivity electrospray ionization-Fourier transform ion cyclotron resonance mass spettrometry of proteins", *Anal. Chem.* **72,** 2271 (2000).
16. S. Berkenkamp, C. Menzel, M. Karas and F. Hillenkamp, "Performance of infrared matrix-assisted laser desorption/ionization mass spectrometry with lasers emitting in the 3 mm wavelength range", *Rapid Commun. Mass Spectrom*. **11,** 1399 (1997).
17. J. Godovac-Zimmerman and L.R. Brown. "Perspective for mass spectrometry and functional proteomics", *Mass Spectrom. Rev.* **20,** 1 (2001).
18. R. Hoffmann, S. Metzger, B. Spengler and L. Otvos, Jr, Sequencing of peptides phosphorylated on serines and threonines by post-source decay in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry", *J. Mass Spectrom.* **34,** 1195 (1999).
19. T. Kinumi, H. Niwa and H. Matsumoto, "Phosphopeptide sequencing by in-source decay spectrum in delayed extraction matrix-assisted laser desorption ionization time-of-flight mass spectrometry", *Anal. Biochem.* **277,** 177 (2000).
20. G. Neubauer and M. Mann, "Mapping of phosphorylation sites of gel-isolated proteins by nanoelectrospray tandem mass spectrometry: potentials and limitations", *Anal. Chem.* **71,** 235 (1999).
21. J.R. Yates, III, "Mass spectrometry. From genomics to proteomics", *Trends Genet*. **16,** 5 (2000).
22. P. Petrilli, P. Pucci, H.R. Morris and F. Addeo, "Assignment of phosphorylation sites in buffalo beta-casein by fast atom bombardment mass spectrometry", *Biochem. Biophys. Res. Commun.* **15,** 28 (1986)
23. J.W. Jarvik and C.A. Telmer, "Epitope tagging", *Annu. Rev. Genet*. **32,** 601 (1998).
24. M.B. Goshe, P.T. Conrads, E.A. Panisko, N.H. Angell, T.D. Veenstra and R.D. Smith, "Phosphoprotein Isotope-Coded Affinity Tag Approach for Isolating and Quantitating Phosphopeptides in Proteome-Wide Analyses", *Anal. Chem*. **73,** 2578 (2001).

25. M.B. Goshe, T.D. Veenstra, E.A. Panisko, T.P. Conrads, N.H. Angell and R.D. Smith, "Phosphoprotein Isotope-Coded Affinity Tags: Application to the Enrichment and Identification of Low-Abundance Phosphoproteins", *Anal. Chem.* **74,** 607 (2002).
26. Y. Oda, T. Nagasu and B.T. Chait, "Enrichment analysis of phosphorylated proteins as a tool for probing the Phosphoproteome", *Nature Biotech*. **19,** 379 (2002).
27. M.B. Smolka, H. Zhou, S. Purkayastha and R. Aebersold, "Optimization of the Isotope-Coded Affinity Tag-Labelling Procedure for Quantitative Proteome Analysis", *Anal. Biochem.* **297,** 25 (2001).
28. S.P. Gygi, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb and R. Aebersold, "Quantitative analysis of complex protein mixtures using isotope-coded affinity tags", *Nature Biotechnol*. **17,** 994 (1999).
29. L. Wells, K. Vosseller, R.N. Cole, J.M. Cronshaw, M.J. Matunis and G.W. Hart, "Mapping sites of O-GlcNAc modifications using affinity tags for serine and threonine post-translational modifications", *Mol. Cell. Proteomics* **1,** 791 (2002).
30. H.L. Goering, D.I. Relyea and D.W. Larsen, "Preparation of thiols: formation from alkenes", *J. Am. Chem. Soc.* **78,** 348 (1956).
31. T.P. Conrads, H.J. Issaq and T.D. Veenstra, "New tools for quantitative phosphoproteome analysis", *Biochem. Biophys. Res. Commun.* **25,** 885 (2002).
32. W. Li, R.A. Boykins, P.S. Backlund, G. Wang and H.C. Chen, "Identification of phosphoserine and phosphothreonine as cysteic acid and beta-methylcysteic acid residues in peptides by tandem mass spectrometric sequencing", *Anal. Chem*. **15,** 5701 (2002).
33. T. Mega, N. Nakamura and T. Ikenaka, "Modifications of substituted seryl and threonyl residues in phosphopeptides and a polysialoglycoprotein by beta-elimination and nucleophile additions", *J. Biochem.* (*Tokyo*) **107,** 68 (1990).
34. H. Jaffe, Veeranna and H.C. Pant, Characterization of serine and threonine phosphorylation sites in beta-elimination/ ethanethiol addition-modified proteins by electrospray tandem mass spectrometry and database searching. *Biochemistry* **17,** 16211 (1998).
35. M. Mann, " Quantitative proteomics?" *Nat. Biotechnol*. **17,** 954 (1999).
36. Y. Oda, K. Huang, F.R. Cross, D. Cowburn and B.T. Chait, "Accurate quantitation of protein expression and site-specific phosphorylation", *Proc. Natl. Acad. Sci. USA* **96,** 6591 (1999).

Received: 5 March 2003
Revised: 17 September 2003
Accepted: 20 September 2003
Web Publication: 2 December 2003