NOVEL METHODOLOGIES IN QUANTITATIVE PROTEOMICS

Angelo Palmese

Dottorato in Scienze Biotecnologiche – XXIV ciclo
Indirizzo Biotecnologie Industriali e Molecolari
Università di Napoli Federico II

Dottorato in Scienze Biotecnologiche – XXIV ciclo
Indirizzo Biotecnologie Industriali e Molecolari
Università di Napoli Federico II

NOVEL METHODOLOGIES IN QUANTITATIVE PROTEOMICS

Angelo Palmese

Dottorando:	Angelo Palmese
Relatore:	Prof. Gennaro Marino
Co-relatore:	Dott.sa Angela Amoresano
Coordinatore:	Prof. Giovanni Sannia

I nostri sogni e desideri
cambiano il mondo.
(Karl Popper)

INDEX

Riassunto

La proteomica è un potente e sensibile strumento per studiare le proteine nel loro "insieme" e permette, in definitiva, di comprendere nella loro complessità i fenomeni biochimici che avvengono nei processi cellulari [1-2]. Se, parafrasando una definizione di biotecnologia della European Federation of Biotechnology, le cellule o parti di esse vengono specificamente utilizzate per la "produzione di beni e servizi" [3], appare evidente come sempre più la proteomica, insieme con le altre discipline omiche, svolga ormai un ruolo centrale nel rispondere alle richieste delle biotecnologie industriali in termini di miglioramenti ed ottimizzazione dei processi esistenti, di sviluppo di nuovi prodotti per l'industria chimica e farmaceutica, di risposta alle sempre più impellenti esigenze di energia da fonti rinnovabili [4]. In una recente rassegna [3] la proteomica è rappresentata come uno degli ingranaggi centrali che consente di estendere gli approcci della "biologia dei sistemi" alle biotecnologie. Recentemente, in seguito ai notevoli progressi tecnologici nel campo della progettazione e della realizzazione di spettrometri di massa ad elevata risoluzione e sensibilità, sono state sviluppate una serie di metodologie per l'analisi di modifiche post-traduzionali, basate sull'utilizzo di tecniche di spettrometria di massa. Tuttavia esiste la necessità di sviluppare metodologie di analisi quantitativa per sfruttare adeguatamente queste innovazioni strumentali e per rispondere in termini quantitativi ad esigenze poste, ad esempio, dall'implementazione dei processi biotecnologici. Un'esigenza particolarmente sentita si riferisce alla necessità di definire in termini quantitativi alcune modifiche post-traduzionali che, a livello sub-stechiometrico, comportano l'accensione o lo spegnimento di importanti circuiti metabolici. Un approccio che potenzialmente è in grado di rispondere a questo tipo di problematica, e che è stato da me utilizzato durante il mio dottorato di ricerca, è rappresentato da procedure di derivatizzazione chimica selettiva mediante molecole isotopicamente marcate, in combinazione con l'utilizzo di tecniche avanzate di spettrometria di massa. Una molecola adatta a questo tipo di applicazione è iTRAQ [5]. La scelta di iTRAQ è motivata dalla possibilità di utilizzare questa molecola per strategie definite RIGhT (Reporter Ion Generating Tag) [6] per l'analisi selettiva di modifiche post-traduzionali; strategie analitiche di questo tipo sfruttando le caratteristiche di frammentazione in fase gassosa degli iTRAQ-derivati e le potenzialità di uno spettrometro di massa dotato di analizzatore ibrido a triplo quadrupolo e trappola ionica lineare (QqLIT), in grado di effettuare analisi utilizzando una particolare modalità di scansione definita *Precursor Ion Scan* (PIS). Con questa modalità è possibile effettuare un'analisi dei soli ioni precursori che durante frammentazione producono uno specifico ione frammento che. Un'analisi di questo tipo, rispetto ad una classica analisi LC-MS/MS, semplifica notevolmente l'interpretazione dei dati: infatti, a meno di falsi positivi, gli unici spettri MS ed MS/MS che si otterranno dall'analisi, saranno quelli relativi ai peptidi marcati con iTRAQ. In questo modo si può effettuare contemporaneamente l'analisi qualitativa e quantitativa dei campioni in esame, evitando ulteriori passaggi di purificazione o arricchimento delle specie di interesse, che sono fondamentali nelle strategie finora proposte per l'analisi delle suddette modifiche. L'eliminazione del passaggio di arricchimento consente di aumentare la sensibilità della strategia analitica, fattore estremamente importante dato il carattere sub-stechiometrico delle modifiche post-traduzionali che sono state oggetto di studio durante lo svolgimento della tesi di dottorato. In particolare durante il mio dottorato mi sono occupato dello sviluppo di metodologie per l'analisi

qualitativa e quantitativa della nitrazione dei residui di tirosina, della carbonilazione dei residui di lisina e della fosforilazione dei residui di serina e treonina.

Analisi quantitativa dei residui di 3-nitrotirosina

La nitrazione di residui di tirosina (conversione della tirosina in 3-nitrotirosina) è una tra le più comuni modifiche post-traduzionali indotte da stress ossidativo [7]; si tratta di una modifica chimica irreversibile, difficilmente analizzabile mediante approcci di proteomica classica dato il suo carattere sub-stechiometrico. Molti degli studi pubblicati sullo studio di questa modifica sono basati sull'utilizzo della elettroforesi bidimensionale in combinazione con la tecnica del Western blot, sfruttando la disponibilità di anticorpi specifici per i residui di tirosina nitrati [8]; dati i limiti di queste metodologie, risulta però molto difficile ottenere informazioni circa la precisa localizzazione del sito di modifica in una proteina, nonché informazioni quantitative sulla stechiometria della modifica in oggetto. Per superare tali limitazioni è stato quindi necessario progettare una strategia di analisi estremamente specifica e sensibile, che consentisse di ottenere sia informazioni qualitative sia quantitative dei residui in esame. Come già specificato questa strategia si è basata sull'impiego di reattivi iTRAQ. Oltre all'ottimizzazione dei processi chimici di derivatizzazione e marcatura, sono stati effettuati studi sulle caratteristiche di frammentazione in fase gassosa delle molecole utilizzate, mediante l'utilizzo di uno spettrometro di massa dotato di analizzatore ibrido a triplo quadrupolo e trappola ionica lineare. La strategia sviluppata durante la mia tesi di dottorato ha previsto: i) l'idrolisi triptica della miscela proteica in esame; ii) la protezione dei gruppi amminici primari (N-term, Lys) mediante acetilazione; iii) la riduzione dei nitrogruppi ad amminogruppi; iv) la marcatura selettiva degli amminogruppi generati. Data l'irreperibilità in commercio di proteine modello per l'analisi dei residui di 3-nitrotirosina, è stato inoltre necessario mettere a punto una strategia di modifica chimica per l'introduzione di nitro-gruppi sui residui di tirosina; tale strategia ha previsto l'utilizzo di tetranitrometano (TNM) [9] per la nitrazione di tali residui. Per l'ottimizzazione dell'intera strategia è stata scelta come proteina modella l'albumina di siero bovino (BSA), proteina che presenta nella sua struttura primaria 20 residui di Tyr; la verifica della avvenuta nitrazione dei residui di Tyr è stata effettuata mediante analisi di spettrometria di massa MALDI-TOF delle miscele peptidiche ottenute dall'idrolisi enzimatica delle proteine modificate. Dai dati ottenuti in questa prima fase è stato possibile verificare che non tutti i residui di Tyr della proteina sono convertiti a 3-NO_2Tyr. Questo dato è in linea con altri dati presenti in letteratura circa la selettività della nitrazione dei residui di Tyr, selettività che dipende essenzialmente dall'accessibilità al solvente di tali residui e dalle caratteristiche chimico-fisiche dell'intorno amminoacidico delle singole Tyr [10]. Negli esperimenti condotti i residui di Tyr che sono stati identificati come nitrati in seguito alla reazione con TNM sono le Tyr in posizione 137, 331, 400 e 451. La protezione dei gruppi amminici primari è stata effettuata mediante acetilazione; esperimenti preliminari sono stati condotti utilizzando anidride acetica come reattivo acetilante, ma analisi di spettrometria di massa hanno evidenziato come la resa della reazione non fosse quantitativa. Per tale motivo è stato successivamente scelto come reattivo l'acido acetico N-idrossisuccinimmide estere, che ha consentito di ottenere l'acetilazione pressoché quantitativa dei gruppi amminici presenti. Solo in seguito alla reazione di acetilazione è stato possibile effettuare la riduzione dei nitrogruppi delle *o*-nitrotirosine ad amminogruppi, suscettibili alla reazione con iTRAQ. Tale conversione è stata realizzata utilizzando $Na_2S_2O_4$ [11] e anche in questo caso la reazione è stata verificata mediante analisi di spettrometria di massa MALDI-

TOF. Per la reazione di marcatura sono stati utilizzati i reattivi iTRAQ disponibili in commercio, nelle varianti isotopiche comunemente denominate iTRAQ 114 e iTRAQ 117, ottimizzando il protocollo di reazione proposto dalla casa produttrice dei reattivi stessi (Applied Biosystems). Successivamente alla reazione di marcatura, le miscele risultanti sono state riunite in modo da ottenere campioni che contenessero BSA nitrata marcata con iTRAQ 114 e con iTRAQ 117 in rapporto molare 1:1 e 1:4 rispettivamente. Le due miscele (1:1 e 1:4) sono state quindi analizzate mediante LC-MS/MS in modalità *Precursor Ion Scan* per gli ioni reporter tipici dei reattivi iTRAQ (m/z 114 e m/z 117), usando un 4000QTrap accoppiato ad un sistema nanoHPLC; in questo modo è stato possibile identificare i siti di nitrazione ed effettuare l'analisi quantitativa relativa dei soli peptidi nitrati con una singola analisi. Per dimostrare l'applicabilità di tale strategia all'analisi proteomica, è stato costruito un opportuno sistema modello di maggiore complessità, aggiungendo ad un intero estratto proteico di *E. coli*, la BSA nitrata. Successivamente è stato utilizzato il protocollo di analisi precedentemente descritto per l'analisi di questa miscela. La miscela è stata infine analizzate mediante LC-MS/MS come descritto in precedenza; l'interpretazione dei dati ottenuti è stata effettuata utilizzando il software MASCOT per l'identificazione delle proteine nitrate presenti nel campione; sebbene in questo caso la miscela fosse molto più complessa, il risultato ottenuto ha consentito di identificare gli stessi siti di nitrazione della BSA già identificati nel caso dell'analisi della singola proteina. La strategia proposta è stata infine utilizzata per identificare i residui di o-nitritirosina nell'intera frazione proteica del latte bovino. In particolare tale frazione proteica è stata nitrata *in vitro*, come per la BSA, e il grado di nitrazione è stato verificato mediante SDS-PAGE e Western blot per verificare che la percentuale di proteine nitrate nel campione fosse bassa, in modo da riprodurre il più possibile i reali livelli di proteine nitrate presenti nelle cellule.
Anche in questo caso i campioni proteici sono stati sottoposti al protocollo di reazione precedentemente illustrato, e marcati con i reattivi iTRAQ 114 e 117; anche in questo caso sono state preparate miscele composte da campione marcato con iTRAQ 114 e campione marcato con iTRAQ 117 in rapporto molare 1:1 e 1:4 rispettivamente. Le successive analisi mediante LC-MS/MS in modalità PIS, hanno permesso sia di identificare i siti di nitrazione presenti sia di effettuare l'analisi quantitativa relativa dei residui nitrati nelle due miscele in esame.

Marcatura selettiva dei residui carbonilati per l'analisi qualitativa mediante spettrometria di massa bidimensionale e per l'analisi quantitativa

Una strategia di analisi simile a quella ottimizzata per l'analisi della nitrazione, è stata utilizzata per l'analisi dei residui di carbonilazione, generati in seguito ad eventi di ossidazione. La carbonilazione è una modifica post-traduzionale chimica irreversibile; consiste nella aggiunta di un gruppo carbonilico (aldeide o chetone) su particolari residui amminoacidici (lisina, arginina, prolina, treonina, istidina) a causa della presenza di elevati livelli di ROS (*Reactive Oxygen Species*) [12]. Durante il mio dottorato mi sono occupato in particolare dell'analisi di residui di semialdeide α-amminoadipica, generati dall'ossidazione di residui di lisina [12]. La carbonilazione di tali residui è una delle maggiori conseguenze dei fenomeni di stress ossidativo; per tale motivo riuscire ad avere informazioni qualitative e quantitative su tali fenomeni, può risultare di fondamentale importanza nella comprensione dei danni cellulari causati dallo stress ossidativo. Per ottenere tale risultato mi sono servito di strategie di tipo RIGhT mediante l'utilizzo combinato di particolari reattivi per la marcatura selettiva dei residui d'interesse in combinazione con tecniche avanzate di analisi di

spettrometria di massa; in particolare in un primo momento ho messo a punto una metodologia per l'analisi qualitativa dei residui di semialdeide α-amminoadipica, sfruttando la dansil-idrazide, un derivato del dansile presente in commercio specifico per i gruppi carbonilici, e in seguito ho sintetizzato un nuovo derivato di iTRAQ, l'iTRAQ-idrazide, per l'analisi quantitativa di tali residui. La strategia proposta risponde alle necessità di sensibilità e specificità per l'analisi di una modifica sub-stechiometrica, quale appunto la carbonilazione. Le metodologie esistenti, basate essenzialmente su tecniche immunochimiche [13], non hanno la capacità di consentire l'identificazione puntuale dei residui modificati, nonché di effettuarne un'analisi quantitativa.

Data l'irreperibilità in commercio di proteine modello con residui carbonilati, è stato innanzitutto necessario mettere a punto una strategia di modifica chimica per l'introduzione di residui di semialdeide α-amminoadipica in una proteina modello. Per ottenere tale sistema modello sono state provate varie condizioni di ossidazione su varie proteine disponibili in commercio; in seguito a tali prove la proteina scelta è stata la ribonucleasi A bovina, che è stata ossidata mediante trattamento con NaClO in condizioni controllate [14]. Dai dati ottenuti in questa prima fase è stato possibile verificare che non tutti i residui ossidabili della proteina sono convertiti nelle rispettive forme ossidate; questo dato è in linea con altri dati presenti in letteratura circa la selettività della reazione di carbonilazione [15]. Ottenuto il modello proteico carbonilato, mi sono dedicato all'ottimizzazione del processo chimico di marcatura dei residui di interesse con dansil-idrazide, e all'ottimizzazione del metodo di analisi mediante LC-MS/MS. La marcatura dei residui di interesse è stata effettuata in seguito all'idrolisi triptica della miscela proteica in esame; l'analisi selettiva dei soli peptidi modificati è stata effettuata sfruttando le caratteristiche di uno spettrometro di massa ibrido QqLIT. In particolare l'analisi è stata condotta mediante LC-MS/MS e i peptidi marcati sono stati isolati mediante due criteri di selezione basati su due esperimenti successivi di spettrometria di massa tandem: un'analisi precursor ion scan (PIS), seguita da un'analisi in MS^3, in una sorta di analisi di spettrometria di massa bidimensionale. La modalità PIS mi ha consentito di isolare all'interno della miscela in esame i soli ioni precursori che in seguito a frammentazione generano lo ione frammento a m/z 170, tipico dei dansil derivati [6]; i precursori così selezionati vengono poi sottoposti ad un'ulteriore analisi di spettrometria di massa in modalità MS^3, sfruttando la transizione tipica dei dansil derivati m/z 234-170. Mediante l'utilizzo di questa analisi di spettrometria di massa multidimensionale è stato possibile isolare dalla miscela peptidica in esame il solo peptide 1-10 della RNasi A, recante un residuo di lisina carbonilato e marcato con dansil-idrazide (in particolare il residuo di lisina in posizione 7). Per verificare la capacità di tale strategia di individuare selettivamente residui carbonilati in miscele più complesse, 100 μg di RNasi A carbonilata, sono stati aggiunti a 10 mg di un intero estratto proteico di *E. coli*. La miscela proteica ottenuta è stata sottoposta al protocollo di marcatura e analizzata mediante la strategia di LC-MS/MS bidimensionale descritta in precedenza. Anche in questo caso è stato possibile individuare un unico peptide marcato che, in seguito ad interpretazione dei rispettivi spettri MS e MS/MS, è stato attribuito nuovamente al peptide 1-10 della RNasi A, carbonilato e marcato con dansil-idrazide sul residuo di lisina in posizione 7. La mancanza di una molecola di dansil-idrazide marcata isotopicamente, non mi ha permesso di utilizzare questo reattivo per l'analisi quantitativa selettiva dei residui carbonilati. Per ovviare a ciò ho modificato chimicamente la molecola di iTRAQ con l'aggiunta di un residuo di idrazina, per dirigerne la specificità ai gruppi carbonilici. È stato, quindi, necessario ottimizzare le

condizioni di reazione e di purificazione per l'ottenimento della molecola di iTRAQ-idrazide. La molecola è stata ottenuta facendo reagire l'iTRAQ con la molecola di idrazina disponibile in commercio; il prodotto è stato successivamente purificato mediante RP-HPLC e verificato mediante analisi di spettrometria di massa MALDI-TOF. L'iTRAQ-idrazide è stata utilizzato per la marcatura dei residui carbonilati della RNasi A, seguendo il protocollo descritto in precedenza per la dansil-idrazide. La marcatura è stata effettuata utilizzando due varianti isotopiche del reattivo iTRAQ, l'iTRAQ 114 e l'iTRAQ 117, con lo scopo di avere due miscele differenzialmente marcate sui residui carbonilati. I campioni marcati sono stati poi riuniti in modo da ottenere due differenti miscele con rapporto molare 1:1 e 1:4, del campione marcato con iTRAQ-idrazide 114 e di quello marcato con iTRAQ-idrazide 117, rispettivamente. Tali miscele sono state analizzate mediante la strategia LC-MS/MS descritta per l'analisi dei residui nitrati. L'interpretazione degli spettri MS e MS/MS ottenuti, mi ha permesso di individuare il residuo di lisina in posizione 7 sul peptide 1-10 della RNasi A carbonilata come unico residuo modificato; in più, la presenza nello spettro di frammentazione dei segnali degli ioni reporter di iTRAQ, ha reso possibile effettuare l'analisi quantitativa relativa delle due miscele analizzate; i dati relative alle aree sottese dei segnali degli ioni reporter di iTRAQ, hanno consentito di calcolare un rapporto sperimentale di 1.07 $\pm$ 0.03 per la miscela 1:1 e di 3.97 $\pm$ 0.17 per la miscela 1:4. Questa metodologia di analisi qualitativa e quantitativa, rappresenta la prima strategia analitica in grado di consentire contemporaneamente la localizzazione e la quantificazione dei residui carbonilati di una proteina; inoltre ha dimostrato di possedere la specificità tale per consentire l'analisi qualitativa e quantitativa di tali residui su scala proteomica.

Analisi quantitativa selettiva dei residui di pSer e pThr

Durante il mio lavoro di tesi mi sono, inoltre, dedicato allo sviluppo di una metodologia analitica per l'analisi selettiva e quantitativa dei residui di pSer e pThr. La fosforilazione è una modifica post-traduzionale reversibile che agisce come un interruttore nella maggior parte dei meccanismi di regolazione dell'attività enzimatica e di trasduzione del segnale [16]. La fosforilazione riveste, quindi, un ruolo centrale nella regolazione dei fenomeni cellulari e riuscire ad avere informazioni sui residui modificati in un dato evento cellulare e sulla stechiometria di tale modifica, risulta di fondamentale importanza per la comprensione dell'evento stesso. Anche la fosforilazione è una modifica sub-stechiometrica e come tale è molto difficile da studiare, soprattutto se si considera il suo carattere di modifica transiente, al contrario della nitrazione e della carbonilazione. Per tali motivi l'analisi proteomica della fosforilazione è ancora oggi di difficile realizzazione. Oltre alle classiche strategie di analisi che sfruttano le potenzialità dell'elettroforesi bidimensionale in combinazione con tecniche immunochimiche e colorimetriche per l'individuazione di proteine fosforilate [17-18], sono state sviluppate una serie di metodologie che fanno uso della spettrometria di massa e di tecniche cromatografiche di arricchimento, in fosfoproteine e fosfopeptidi, delle miscele in esame [19-20]. Durante il mio lavoro di tesi ho sviluppato una strategia per l'analisi quantitativa selettiva dei residui di pSer e pThr, basata sulla modifica chimica selettiva dei residui di interesse e sull'analisi mediante tecniche avanzate di spettrometria di massa tandem. La strategia prevede la modifica chimica dei residui di pSer/pThr utilizzando un nuovo reattivo di classe RIGhT basato su molecole iTRAQ da me modificate con l'aggiunta di un residuo di cisteammina. In accordo con le cosiddette strategie di proteomica "gel free" la strategia prevede che la marcatura sia effettuata sulla miscela peptidica dopo idrolisi

triptica dei campioni. La procedura consiste in una reazione di β-eliminazione del gruppo fosfato dei residui di pSer/pThr, seguita da addizione di Michael del risultante residuo α,β-insaturo con la molecola di iTRAQ-cisteammina precedentemente sintetizzata. Questa molecola è stata scelta sia perché risponde alla necessità di avere un gruppo SH libero da utilizzare come donatore per la reazione di addizione di Michael, sia perché possiede un residuo iTRAQ, necessario per l'analisi quantitativa selettiva. Per questo motivo il primo obiettivo del mio lavoro è stato quello di sintetizzare e purificare tale molecola. La sintesi è stata messa a punto utilizzando cistammina e iTRAQ come prodotti di partenza; la molecola di iTRAQ-cistammina ottenuta è stata ridotta mediante l'utilizzo di TBP ad iTRAQ-cisteammina e purificata mediante RP-HPLC. Il prodotto della reazione è stato verificato mediante analisi di spettrometria di massa MALDI-TOF. La procedura di marcatura selettiva in due passaggi è stata ottimizzata utilizzando come proteina modella la α-caseina bovina, una proteina naturalmente fosforilata. La reazione di β-eliminazione è stata effettuata incubando un'aliquota di miscela peptidica di α-caseina con una soluzione satura di $Ba(OH)_2$ [6]; è infatti noto che a pH elevati, i residui di pSer/pThr perdono il gruppo fosfato in seguito a β-eliminazione catalizzata da basi [21]. I residui α,β-insaturi generati sono ottimi accettori di Michael e sono stati modificati mediante reazione con iTRAQ-cisteammina; in particolare due differenti aliquote di miscela peptidica β-eliminata sono state sottoposte ad addizione di Michael con due varianti di iTRAQ-cisteammina, sintetizzate a partire rispettivamente dalle varianti 115 e 116 di iTRAQ. I campioni differenzialmente marcati sono stati successivamente riuniti in modo da ottenere tre differenti miscele con rapporti molari 1:1, 1:2 e 1:4 tra campione marcato con il derivato di iTRAQ 115 e quello marcato con il derivato di iTRAQ 116, rispettivamente. Le tre miscele sono state analizzate separatamente mediante LC-MS/MS in modalità precursor ion scan, sfruttando le potenzialità di uno spettrometro ibrido QqLIT accoppiato ad un sistema nanoHPLC. Utilizzando la PIS è stato possibile selezionare all'interno della miscela in esame le sole specie in grado di generare, in seguito a frammentazione, il gruppo reporter tipico delle molecole iTRAQ; di conseguenza le sole specie rivelabili sono quelle in cui il gruppo fosfato è stato sostituito con un residuo di iTRAQ-cisteammina. Con questa analisi è stato possibile isolare selettivamente il peptide 104-119 con un residuo di pSer in posizione 115 modificato con iTRAQ-cisteammina. Inoltre, utilizzando il valore delle aree sottese ai segnali dei gruppi reporter di iTRAQ negli spettri MS/MS, è stato possibile calcolare i rapporti quantitativi per le tre miscele in esame, ottenendo risultati molto vicini ai rapporti teorici: 1,02 $\pm$ 0,08 per la miscela 1:1, 2,06 $\pm$ 0,06 per la miscela 1:2 e 3,99 $\pm$ 0,05 per la miscela 1:4. Per dimostrare l'applicabilità di tale strategia all'analisi proteomica, è stato costruito un opportuno sistema modello di maggiore complessità, aggiungendo ad un intero estratto proteico di *E. coli*, un'aliquota di α-caseina β-eliminata. La miscela è stata quindi sottoposta al protocollo di marcatura e analizzata mediante LC-MS/MS come descritto in precedenza; l'interpretazione dei dati ottenuti è stata effettuata utilizzando il software MASCOT per l'identificazione delle proteine modificate. I risultati ottenuti da tale analisi hanno consentito di identificare selettivamente le sole specie marcate con iTRAQ-cisteammina, consentendo di individuare i siti di fosforilazione e di effettuarne l'analisi quantitativa anche in miscele complesse.

BIBLIOGRAFIA

1. Zhou, M.; *et al*. *Proteomics*. **2007**, *7*, 2688-2697.

2. Josic, D.; *et al. Proteomics*. **2007**, *7*, 3010-3029.
3. Sang, Y.L.; *et al. TRENDS in Biotechnology*. **2005**, *23* (7), 349-358.
4. Hermann, T. *Curr Opin Biotechnol*. **2004**, *15* (5), 444-448.
5. Ross, P.L.; *et al. Mol Cell Proteomics*. **2004**, *3*, 1154-1169.
6. Amoresano, A.; *et al. Rapid Commun Mass Spectrom*. **2006**, *20* (9), 1400-1404.
7. Schopfer, F.J.; *et al. Trends Biochem Sci*. **2003**, *28* (12), 646-654.
8. Butt, Y.K.; *et al. Methods Enzymol*. **2008**, *440*, 17-31.
9. Amoresano, A.; *et al. Anal Chem*. **2007**, *79*, 2109-2117.
10. Abello, N.; *et al. J Proteome Res*. **2009**, *8* (7), 3222-3238.
11. Sokolovsky, M.; *et al. Biochemistry*. **1966**, *5*, 3582-3589.
12. Requena, J.R.; *et al. Proc Natl Acad Sci USA*. **2001**, *98* (1), 69-74.
13. Choi, J.; *et al. Free Radic Biol Med*. **2004**, *36*, 1155-1162.
14. Headlam, H.A.; *et al. Free Radic Biol Med*. **2004**, *36*, 1175-1184.
15. Nystrom, T.; *et al. Embo J*. **2005**, *24*, 1311–1317.
16. Cohen, P. *Trends Biochem Sci*. **2000**, *25*, 596-601.
17. Baik, J.Y.; *et al. J Biotechnol*. **2008**, *133*, 461-468.
18. Masaki, S.; *et al. Biotechnol Lett*. **2008**, *30* (5), 955-960.
19. Wang, J.L.; *et al. Proteomics*. **2006**, *6*, 404-411.
20. Pinkse, M.W.; *et al. J Proteome Res*. **2008**, *7*, 687-697.
21. Conrads, T.P.; *et al. Biochem Biophys Res Commun*. **2002**, *25*, 885.

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 and ii) diversity of protein expression (multiple protein forms).
Post-translational modification of proteins is part of what makes proteomics so much challenging. It is an event with dramatic effects on the complexity of the proteome. Nowadays the attention towards PTMs is notably increased due to the fact that they play a critical role in cellular functioning and they vary in response of environmental stimuli. There are many classical techniques designed to analyse PTMs.
This PhD thesis focuses on the development of new proteomics approaches for qualitative and quantitative analysis of protein nitration, protein carbonylation and protein phosphorylation, three of the most important known PTMs. The approaches are mainly based on selective chemical modification of residues of interest coupled to advanced mass spectrometry methods. These approaches aim at a comprehensive analysing of such modifications by identification of the modified proteins, exact localization of the modified residues and quantification of the modification.
The first section of this thesis is focused on the development and the optimization of a new strategy for the qualitative and quantitative analysis of nitration of tyrosine residues. Thus, improved approach, based on chemical labelling and mass spectrometric procedures, for the selective analysis of 3-nitrotyrosine residues in protein mixtures is described. However the major aim of this work is devoted to the use of isotopically labelled molecule to allow a simple and direct quantitative MS analysis.
The second part aims to the optimization of a labelling protocol of carbonylated residues coupled to advanced mass spectrometry based scanning techniques to improve detection of such residues. Moreover a method for selective quantitative analysis is also described, aimed at the direct quantification of oxidation levels of proteins in different samples.
The last part of the work is focused on the development of a strategy for qualitative and quantitative analysis of serine and threonine phosphorylation without preliminary enrichment but using chemical labelling in combination with the high performance of a novel hybrid mass spectrometer Linear Ion Trap.
To prove their wide enforceability these methodologies were set up first on model proteins and then applied to more complex systems.

I. INTRODUCTION

I.1 Proteomics in modern Biotechnology

The relevance of biotechnology in the production of industrial biochemicals is growing steadily owing to the development of new products, new production processes and, in the future, the growing need to use renewable resources. Major biotechnological products include bioethanol, amino acids, polymers, vitamins and enzymes. The development of powerful analytical technologies, usually summarized by the term "*functional genomics*", has a significant role in this growth. These techniques include genomics, transcriptomics, proteomics, and metabolomics together with carbon flux determinations. Most of these technologies, developed over the past decades, are more or less strongly dependent on the deciphering of genomes. Nevertheless, even if the knowledge of an organism's genome is necessary, the "*omics*" era has shown that it alone is not sufficient to understand the life of an organism. On these bases, transcriptomics and proteomics have helped us to better understand how these organisms transfer their genetic information into life. Even though proteomics technologies were first developed more than twenty years ago, this technique has become a powerful tool with the availability of genetic information obtained from genome sequencing projects. In fact in the last few years proteomics has become a powerful tool for the discovery of new proteins and the investigation of complex biochemical processes and protein-protein interactions [1-2]. In medical sciences, proteomics techniques lead to the discovery of disease biomarkers, especially in cancer and neurodegenerative diseases, becoming key milestones on the way to better and more targeted medical treatment [3-5]. Proteomics methods are mostly based on mass spectrometry, a relative expensive tool that at the beginning limited the spread of proteomics in food technology and biotechnology. During the last years only few works have been presented really dealing with the application of proteomics in biotechnology [6-7]. However, it has changed rapidly and proteomics technologies are becoming more practical, and the terms "industrial process proteomics" [8] and "industrial proteomics" [9] have already been used. Recently in a review by Gupta and Lee [10] it is described the use of genomic and proteomic techniques in the design and optimization of bioprocesses, starting with cell culture and fermentation, and continuing with further process development steps such as downstream processing, scale up, and process characterization and validation. An example of efforts made to improve proteomics technologies, can be found in the production of amino acids by using *C. glutamicum* cells [11]. Proteomics studies of *C. glutamicum* have been reviewed [12] and the relatively new field of post-translational modifications has been described [13]. Similar works describe applications of these technologies to other industrially relevant organisms, such as *B. subtilis*, *C. acetobutylicum* and *E. coli* [14-19]. Nowadays proteomics has finally progressed into different branches of biotechnology. It is also shown that proteomics technology can be very useful in development of production processes for therapeutic proteins by use of genetically engineered animal cells [10,20] or human stem cells [21]. In "classical" fermentation industry, proteomic methods can also be used for identification of targets for bioprocess improvement [22] and for quality control of industrial processes of other food products of animal origin [23]. Unlike mammalian cells that are used almost exclusively for production of protein-based therapeutics [10], microbial cells are also widely used for production of small molecules [22]. Microbial fermentation in food

processing has been a tradition for thousands of years [24]. The genomes of most economically important microbial cells are already sequenced and proteomic technologies can be applied during various process development steps, starting with the selection and optimization of the functions of the industrial strains, application of the knowledge of cell function in response to the changes of production parameters, validation of the downstream processing, and thorough characterization of the final product (figure I.1).

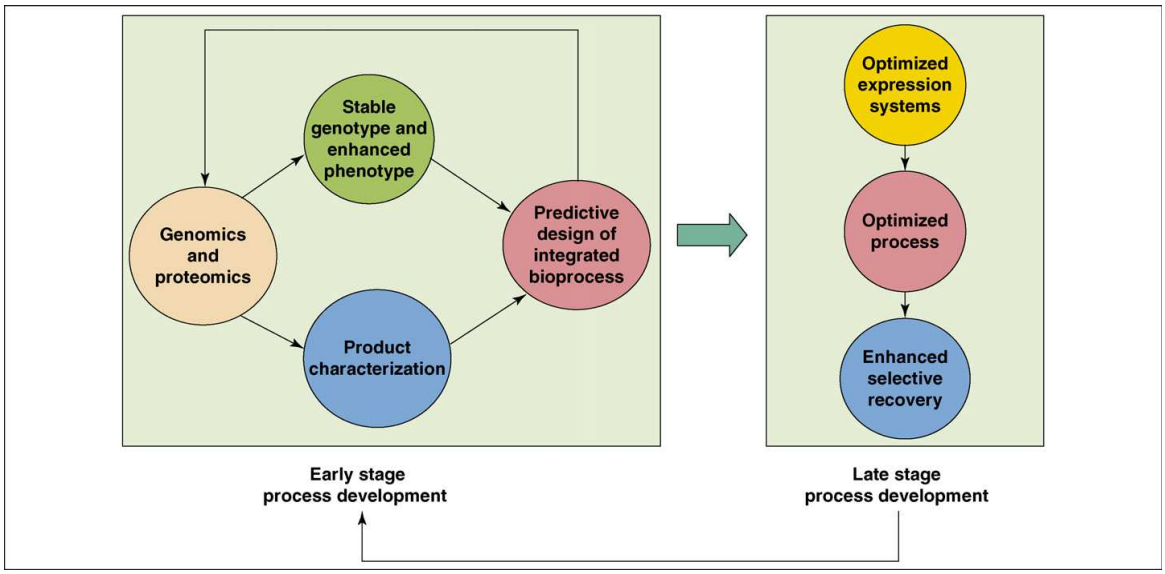


Figure I.1. *A possible role of genomics and proteomics tools in process development. Fundamental understanding of genomic and proteomic processes can facilitate the selection of productive organisms. It can also facilitate product or impurity characterization between different cell lines, leading to better integration of upstream and downstream processing operations. These resulting systems can then be used at late-stage process development to validate and scale-up the process.*

I.2 Proteomics analysis

In 1994, Wilkins defined the Proteome as the "protein complement to a genome" [25]. Proteomics is the study of the expressed protein complement of a genome at a specific time [26]. The terms "proteomics" and "proteome" mirror the terms "genomics" and "genome". Proteomics was only introduced as a concept in the mid 1990s but it has deep roots in the early 1970s, as a result of important technical advances in biochemistry and genetics. The drivers of genomic and proteomic analyses are the technological achievements of the past decade that enable the quantitative analysis of the DNA sequence, mRNA and protein expression inside cells and include tools such as DNA microarrays, two-dimensional gel electrophoresis (2D-GE) and mass spectrometry (MS). The application of these methods in drug discovery has heralded a new era for target identification [27-30]. These genome-wide approaches have been adopted by the biotechnology and pharmaceutical industries to complement traditional approaches for target identification and validation, for hypotheses generation and for experimental analyses in traditional-based methods.
Proteomics arose directly from protein chemistry, taking advantage of analytical tools of the latter, but moving away from the latter. In fact proteomics is based on the idea that to really understand the cellular function and its molecular mechanisms, in which proteins play a key role, it is necessary to study the complete proteome as a single "complex analyte" rather than as a "mixture" of isolated proteins.

For this reason proteomics is a core part of *functional genomics*, which relies on high-throughput approaches for determining the relations between a given genome and the corresponding phenotypes.
Proteome analysis presents specialized analytical problems in two major areas:

- *expression proteomics*;
- *functional proteomics*.

Expression proteomics studies the variations in expression levels that occur at some proteins in a given tissue following specific stimuli. This approach aims to compare patterns of protein expression in a given tissue under certain conditions. The great advantage of expression proteomics upon transcriptomics is that it studies and measures directly the final protein product rather than the levels of mRNA. In this way it is possible to evaluate quantitative and qualitative differences among proteins in samples. Moreover expression proteomics aims to define profile patterns of PTMs of proteins [31]. However it is not enough to map and identify all the proteins that are extant in cells, they must also be assigned with respect to their function and interactions with other proteins or genes. This means that the proteome must be studied under *in vivo* conditions and that the dynamic properties of the proteome must be explicitly exploited to obtain such information. This is the target of functional proteomics that can be defined as the use of proteomics methods to monitor and analyze the spatial and temporal properties of the molecular networks and fluxes involved in living cells [32-33], often using mathematical modelling or other systems biology tools. It also provides readouts for chemical biology and chemical genetics that are necessary to interpret the action of drugs.
Proteomics have to deal with several analytical problems. First of all protein samples presents a high complexity in terms of number of analytes; moreover these analytes could have a very huge range of concentration in a sample that reflect their concentration in a cell or in a tissue [34]. An additional layer of complexity is possible for the presence of different isoforms of proteins that arise not only by differential splicing of mRNA but mainly as a result of the plethora of PTMs that are possible, such as phosphorylation, nitration, acetylation, glycosylation and so on. Moreover PTMs are often of low stoichiometry and transient nature. For example phosphorylation stoichiometry was found between < 10% and 90% depending on the phosphorylation site and protein, but is rarely being determined on a proteome-wide scale [35]. Therefore analysis of the proteome requires a huge set of tools that fulfil the necessity of scientist for sensitivity, specificity, high dynamic range and that are capable of high throughput applications.
Nowadays proteomics can rely on powerful analytical protein-separation technologies (chromatography, electrophoresis), that serve to simplify complex protein mixtures to compare apparent differences in protein levels between two samples. However the most powerful analytical tools for proteomic analysis is mass spectrometry (MS), whose instrumentation has undergone huge changes over the past years, culminating in the development of highly sensitive, robust instruments that can reliably analyze biomolecules, particularly proteins and peptides.

I.2.1 Mass Spectrometry in Proteomics

Mass spectrometry plays a key role in proteomics studies because of its capability of identifying proteins and characterizing post-translational modifications. Most

important features of mass spectrometry are: reproducibility, sensitivity, resolution and accuracy. In particular mass spectrometry is a reproducible methodology based on the determination of the molecular mass, which is not dependent on the experimental conditions.

Principles and instrumentations

Mass spectrometry is based on the production of gas phase ions [36]. These ions can interact with an electric field and can be resolved following their electro-dynamic attitude, which is dependent on their mass-to-charge ratio. In principle all molecular and atomic ions are accessible to mass spectrometric analysis; for this reason mass spectrometry is a very powerful tool for chemical and biochemical analysis in general. Today a wide variety of mass spectrometers are available, all sharing the capability to assign mass-to-charge values to ions, although the principles of operation and the types of experiments that can be done on these instruments differ greatly.
Mass spectrometers have four essential parts: a) a system for sample introduction b) a source that produces gas phase ions from the sample c) one or more mass analyser to separate ions d) an ion detector.
Sample under investigation has to be introduced into the ion source of the instrument, where molecules are ionized. Then these ions are extracted into the analyser region where they are separated according to their mass-to-charge ratios (m/z). The separated ions are detected and the generated signals are collected and sent to a data system where the m/z ratios and their relative abundances are presented in form of an m/z spectrum (figure I.2).

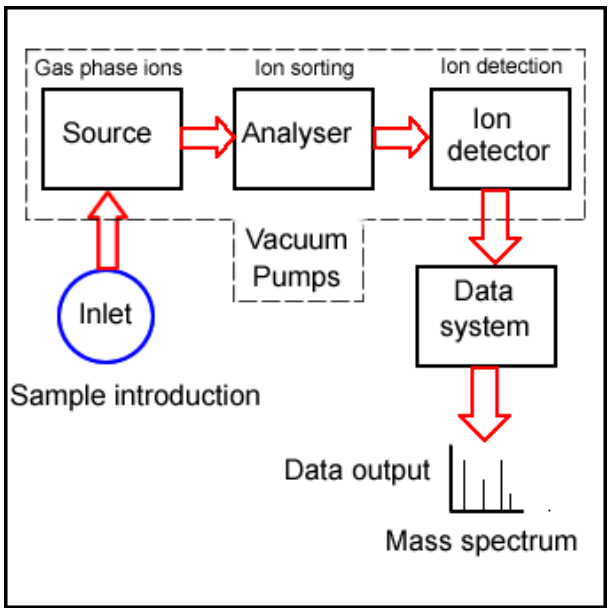


Figure I.2. *Process of measuring of a mass spectrometer.*

Generally mass spectrometers can be distinguished both on the type of ionization systems and on the type of analyser; the latest one is an essential component to define the accessible mass range, sensitivity and resolution of an instrument. All mass spectrometers are operated at very low pressure (from 10^{-5} to 10^{-9}) in order to prevent collisions of ions with residual gas molecules in the analyser during the flight from the ion source to the detector.
In the 1970s mass spectrometry was limited to the analysis of volatile and low-molecular-weight components, because of the absence of ion sources able to transfer organic macromolecules (such as peptides, proteins, nucleic acids, lipids and so on) from liquid or solid state to gaseous phase. These problems were solved with the development of new ionization methods and in particular with the introduction of Electrospray Ionization (*ESI*) [37] and Matrix Assisted Laser Desorption Ionization (*MALDI*) [38] techniques, that are the most common ionization techniques used for biological applications. These ionization techniques are considered "*soft*" ionisation methods, in the sense that the generated ions undergo poor fragmentation.

For MALDI ion source, analytes needs to be embedded in a solid matrix which absorbs energy at the wavelength of the laser. A pulsed laser hits clusters of analytes and matrix, leading to intense heating and generating a plume of ejected material that rapidly expands and undergoes cooling, in a "*pulsed beam*" fashion. Generation of ions is believed to arise through ion/molecule reactions in the gas phase. Generally, $[M+H]^+$ ions are preferentially formed in the positive ion mode, and $[M-H]^-$ ions in the negative ion mode. The presence of singly-charged signals in MALDI spectra allows to correlate each peak one-to-one with analytes, thus allowing the analysis of complex mixtures.
ESI starts by samples in aqueous solution. Analytes exist as ions in solution because they contain functional groups whose ionization is controlled by the pH of the solution. Samples are dissolved in a polar, volatile solvent and pumped through a narrow, stainless steel capillary. A high voltage is applied to the tip of the capillary and the sample emerging from the tip is dispersed into a spray of highly charged droplets; this process is aided by a co-axially introduced gas, called *nebulizing gas*, flowing around the outside of the capillary. The charged droplets diminish in volume by solvent evaporation, assisted by a warm flow of nitrogen, known as *desolvatation gas*, which flows perpendicularly to the front of the ionization source. Once the droplets have reached the Rayleigh limit, ions are desorbed from the droplet generating gas-phase ions in a "*continuous beam*" fashion. This technique gives rise to multiply-charged molecular-related ions such as $(M+nH)^{n+}$ in positive ionization mode and $(M-nH)^{n-}$ in negative ionization mode; the number of charges depends on the chemical-physical characteristics of analytes; thus each analytes could assume more than one charge complicating the form assumed by the signal. For this reason this source is not useful to analyse complex mixtures except using pre-fractionation systems. In fact most uses of ESI-MS are made in combination with chromatographic separation, the so called LC-MS technique. A great enhancement of ESI ion sources has come from the reduction of the flow rate of the liquid used to create the spray to a *nano-scale* level. This device leads to a higher efficiency in creating ions [39] because the charge density at the Rayleigh limit increases significantly with decreasing droplet size (z_R α $8R^{3/2}$). Others advantages related to the use on μLC and nLC are the low consumption of sample and the higher sensitivity of the methods because of the increase in the concentration of the analyte as it elutes off the column.
Once generated, gas phase ions must be analysed/separated; usually mass analysers act as mass filters, based on the interactions of charged analytes with electric and magnetic fields. The revolution in biological mass spectrometry was not only due to improvements in ionization techniques but also to improvements in mass analysers such as *quadrupoles* (*Q*), *ion traps* (*IT*), *time-of-flight* (*TOF*), *Fourier transform ion cyclotron resonance (FTICR)*, *Orbitrap* or combination of these in "*hybrid instruments*", that are commonly used in proteomics (figure I.3); they differ for their resolution, sensitivity and accuracy and for the types of experiments that are able to perform. TOF analysers are typically used in combination with MALDI (MALDI-TOF MS instruments); however both MALDI sources and TOF analysers can be used in different configurations.
In a TOF analyser (working under high vacuum (10^{-8}-10^{-9} atm)) ions can be separated during their flight in a field-free drift region basing upon the time they use to pass through the tube. Their different velocities depend on their m/z ratios. All ions entering the TOF tube have a fixed kinetic energy which is proportional to the applied accelerating voltage and the charge. This implicates that the higher is the mass of

the ion the lower is its velocity and the longer it takes before the ion reaches the detector. A detector, which has been triggered by the laser pulse, records the time-of-flight of the ions. MALDI ionization process causes a spread of kinetic energy of ions resulting in different points, in time and space, of ions formation within the ion source. Thereby ions with the same mass obtain different kinetic energies and velocities during their extraction out of the ion source. This results in peak broadening, causing a loss in resolution. This peak broadening has been reduced by the introduction of the "*delayed extraction*" and of a "*reflectron*" at the end of the linear flight tube [40]. The reflectron allows ions to slow down and reverse their flight path to a second detector. Ions with lower kinetic energy do not penetrate the reflectron as deep and thus turn around faster, catching up with ions of slightly greater kinetic energy that penetrate the reflectron deeper. Thereby the flight times of ions with identical m/z values, but different kinetic energy values, will be corrected when the ions arrive to the detector. TOF based instruments have high resolution and accuracy characteristics.

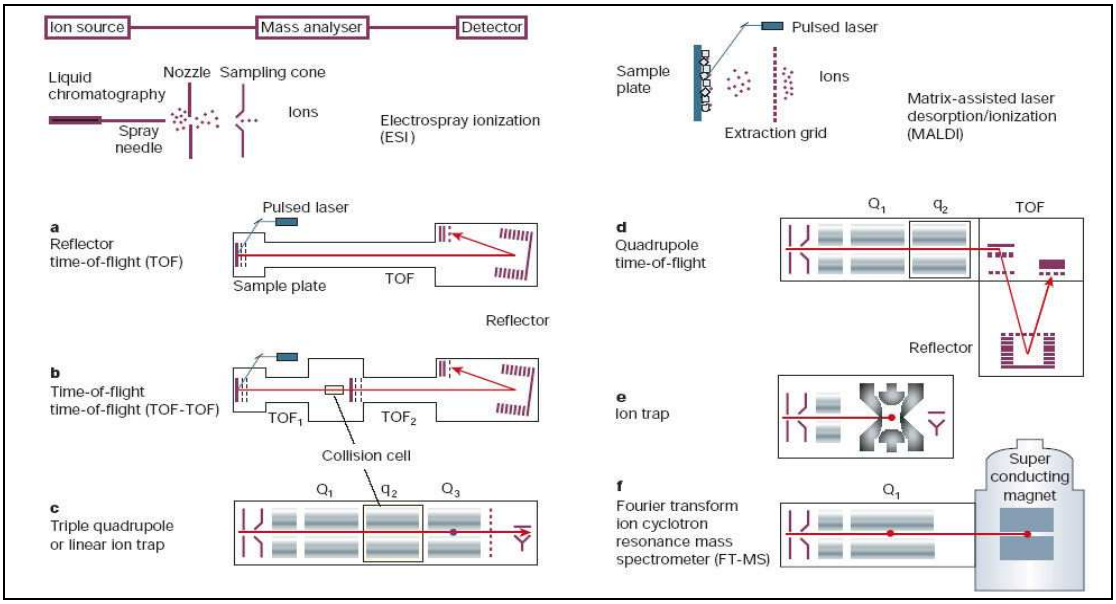


Figure I.3. *Typical mass spectrometers used for proteomics analysis.*

Quadrupole (Q) mass analysers have become one of the most widely used types of mass analyser mainly because of their ease of use, small size and relative low cost. A quadrupole mass filter [41] consists of four parallel rods to which fixed DC and alternating RF potentials are applied. Ions drift through the middle of the quadrupole to the mass detector. Ions oscillate in the x,y-plane and the frequency of oscillations depends on their m/z values and on the applied potentials. If the oscillation of an ion is stable, the ion will pass through the rods and reach the detector. All other ions do not have a stable trajectory through the quadrupole mass analyzer and will collide with the quadrupole rods, never reaching the detector. The m/z values of the ions able to move on stable trajectories depend on the amplitude of RF and DC voltages. The RF is varied to bring ions of different m/z into focus on the detector and thus build up a mass spectrum.

The quadrupole ion trap (QIT) is based on the same principle as the quadrupole mass filter, except that the quadrupole field is generated within a three-dimensional trap. The trap consist of three electrodes, a rotationally symmetrical ring electrode of

hyperbolic shape and two hemispherical end cap electrodes of the same cross-section. One end cap electrode usually has a hole in it to allow ions to enter the trap. Ions are injected axially and to overcome the exit of the ions, their high kinetic energy is dissipated by cooling them collisionally, filling the trap with gas molecules [42]. Ions are dynamically stored in a three-dimensional quadrupole ion storage device by an electric quadrupole field generated by the applied RF voltage. The RF voltage can be scanned to eject successive mass-to-charge ratios from the trap to the detector. The advantages of the ion-trap mass spectrometer include compact size and the ability to trap and accumulate ions (typically the ion trap can hold up about 10^5–10^6 ions before coulombic repulsions significantly affect their trajectories and reduce the mass resolution) to increase the signal-to-noise ratio of a measurement. Mass spectrometer equipped with QIT technology has good sensitivity but low resolution and mass accuracy. Due to its small trapping volume, QIT has a limited trap capacity; when the number of ions in the trap is too high, we assist to deterioration in the mass spectrum and loss of dynamic range due to space-charging interactions. To avoid these effects, the number of ions introduced into the trap has to be strictly controlled [43]. Trapping of ions can also be performed in linear ion trap devices (LIT) [44-45]. LITs have two major advantages over QIT: a larger ion storage capacity and a higher trapping efficiency. In LIT devices analysis is performed by ejecting the ions radially, through the slits of the quadrupole rods, using the mass instability mode as a QIT. Detection is performed by two detectors placed axially along to the rods [46].
An *FTICR* mass spectrometer is essentially composed by an ICR cell, which typically consists of pairs of plates arranged as a cube or capped cylinder, and a surrounding strong magnetic field (4,7-13 Tesla) [47], so that one of the opposing pair of plates is orthogonal to the direction of the magnetic field. Ions are trapped in the cell by an electric field applied to the trapping electrodes and by the magnetic field. Trapped ions move on a circular orbit (cyclotron motion) with orbital frequencies inversely proportional to their mass-to-charge ratios. An RF sweep is then applied across the excitation plates. Ions with cyclotron frequencies in resonance with the applied RF are accelerated and increase their orbit radius bringing them closer to detector plates, which then record the image current. The complex image current is deconvoluted by fast Fourier transformation from which the mass-to-charge ratio is then determined. FTICR mass spectrometers are the most accurate and the highest resolving power instruments currently available [48].

Tandem Mass Spectrometry

Mass spectrometric analysis produces ions that are separated by m/z and analysed directly. Using a soft ionization method the mass spectrum will yield the molecular weight values of single analytes but little or no structural information. Tandem mass spectrometry (MS/MS) is used to produce structural information about a compound by fragmenting specific analytes inside the mass spectrometer and analysing the fragment ions. Tandem mass spectrometry also enables specific compounds to be detected in complex mixtures on account of their specific and characteristic fragmentation patterns. MS/MS is based on two stages of mass analysis, one to select a precursor or *parent* ion and the second to analyze fragment or *daughter* ions, using two different mass analysers; the first mass analyser is used to selectively pass an ion into another reaction region while the second mass analyzer is used to record the m/z values of the dissociation products. Generally the reaction region, or collision cell, is the region into which an inert gas (e.g. nitrogen, argon, helium) is

admitted to collide with the selected ions causing the fragmentation. This process is known as *Collisionally Induced Dissociation* (CID) or *Collisionally Activated Dissociation* (CAD) [49]. Inelastic collisions between the precursor ion with a high translational energy and a neutral target gas, cause the conversion of part of the translational energy into internal energy of the ion, leading to subsequent decomposition [50]. The CID process is highly dependent on the relative masses of the two species. The overall CID process is assumed to occur by a two-step mechanism, where the excitation of the precursors and their fragmentations are separated in time. Generally fragmentation of a precursor ion can occur if the collision energy is high enough to allow the ion to be excited beyond its threshold of dissociation. All CID processes can be separated into two categories based on the energy of the precursor ion. Low-energy collisions are common in triple quadrupoles (QqQ), hybrid devices, such as QqLIT and QqTOF, trapping devices, such as ion traps (IT), and Fourier-transform ion cyclotron resonance (FTICR) instruments; they occur in the 1–100 eV range of collision energy and, even though the average energy deposited per collision may be lower than in high-energy CID, product ion yields are very high, especially because multiple collisions are allowed, for example, by the time chosen for CID in ion trap devices or by changing the gas pressure in a QqQ. High-energy collisions, seen in sector and TOF/TOF instruments, are in the keV collision energy range; the precursor ion beam, having a kinetic energy of a few keV, can enter the collision cell, usually causing single collisions before mass analysis of the product ions. Moreover high-energy CID spectra usually show increased side-chain fragmentation [51]. The observed fragmentation pattern depends on various parameters including the amino acid composition and size of the peptide, excitation method, time scale of the instrument, the charge state of the ion, etc [52]. Peptide precursor ions, dissociated by the most usual low-energy collision conditions, fragment along the backbone at the amide bonds [53-55] forming structurally informative sequence ions and less useful non-sequence ions by losing small neutrals like water, ammonia, etc. The amino acid backbone has three different types of bonds (NH-C_αH, C_αH-CO and CO-NH) and each of them can be fragmented originating different fragment ions. Each bond breakage gives rise to two species, one neutral and the other one charged, and only the charged species is monitored by the mass spectrometer. Hence there are six possible fragment ions for each amino acid residue and these are labelled as in the figure I.4, with the a, b, and c ions having the charge retained on the N-terminal fragment, and the x, y and z ions having the charge on the C-terminal fragment. The most common cleavage sites are at the CO-NH bonds which give rise to the b and/or the y ions [56]. The most comprehensive model currently available to describe how protonated peptides fragment is the so-called "*mobile proton*" model [52]. According to this model peptides activated under low-energy CID, fragment mainly by charge directed reaction. In fact protonation of the amide nitrogen along the amino acid backbone leads to considerable weakening of the amide bond and leads to cleavage, generating fragment ions. 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 this case the "*heterogeneous population model*" [57] assumes that, when the second proton is mobilized along the amino acid backbone, it will exist different protonated forms of the precursor ion that can easily fragment giving complementary series.

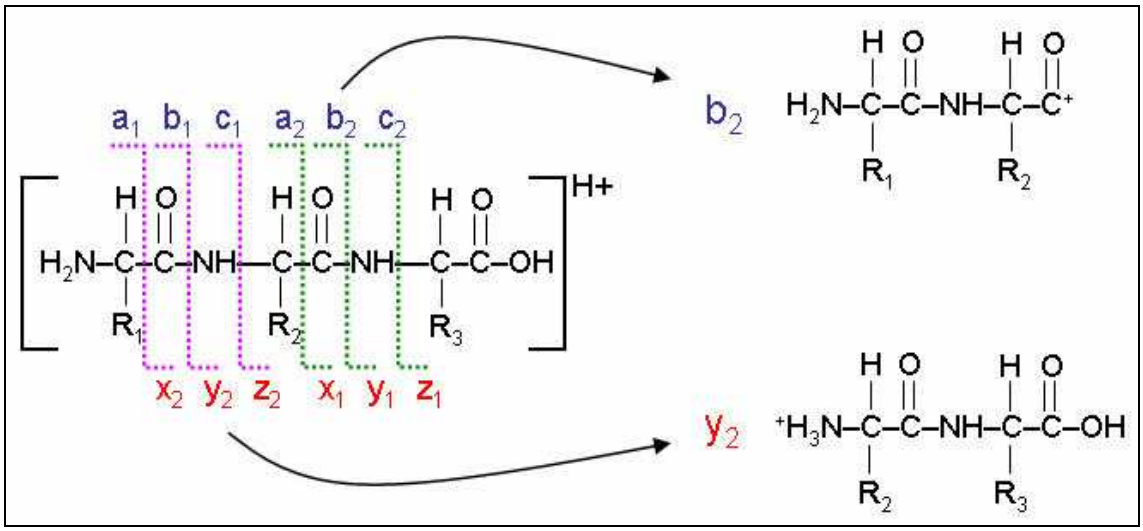


Figure I.4. *Peptides fragmentation scheme.*

Instruments for tandem mass spectrometry can be classified as *tandem in space* or *tandem in time* (figure I.5).

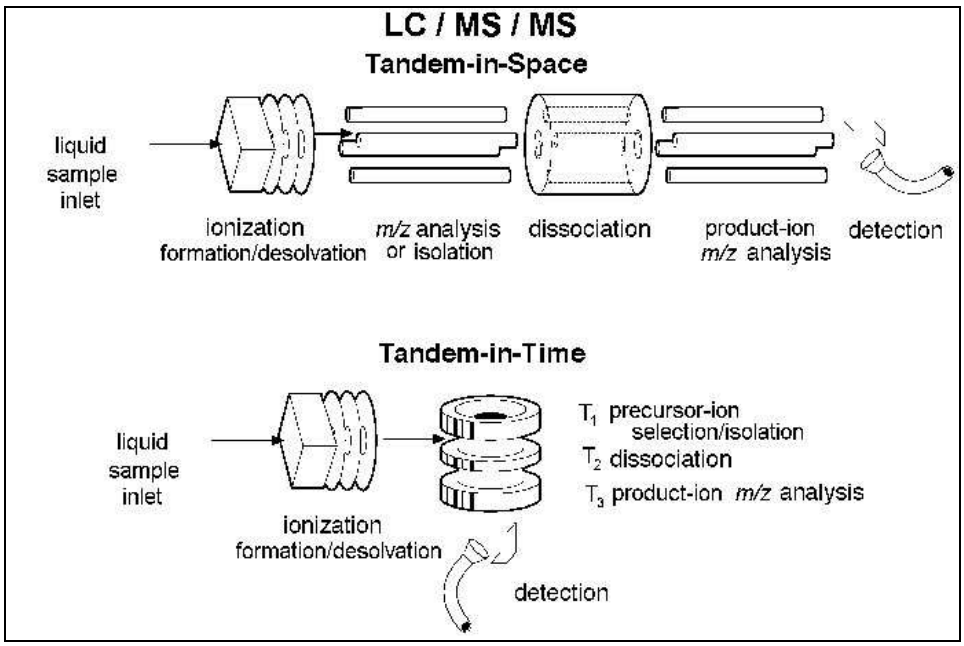


Figure I.5. *Schematic representation of tandem-in-space and tandem-in-time instruments.*

Tandem in space means that ion selection, ion fragmentation and fragments analysis, are events that occur in three different regions of the spectrometer; instruments of this type are the triple quadrupole (QqQ) and hybrid instruments such as QqTOF, QqLIT or TOFTOF. In tandem in time instruments, those three steps of analysis occur in the same region of the spectrometer but in different times; QIT and LIT are instruments classified as tandem in time.

For triple quadrupole mass spectrometers several scan modes are available. The most common scan is the *product ion scan*, where a specific precursor is selected in Q_1, fragmented in q_2 and the products are subsequently analysed in Q_3. A second scan is the *precursor ion scan* (PIS), in which the Q_3 is set to select a specific product

ion formed in q_2 and the Q_1 is scanned for all precursor ions generating the fragment of interest. A third experiment is the *neutral loss scan* (NL) in which both the Q_1 and Q_3 scan and collect data across the whole m/z range, but the two are off-set so that the second analyser allows only those ions which differ by a certain number of mass units (equivalent to a neutral fragment) from the ions transmitted through the first analyser. Multiple stages of CID can be performed in IT instruments, although the experiment is restricted to the *product ion scan*. Several hybrid geometries have recently been introduced with specific advantages, including the quadrupole-linear ion trap (QqLIT) and the quadrupole-time of flight (QqTOF) instruments. The former allows triple quadrupole scan types as well as MS^3 capabilities in the linear ion trap [58], while the latter yields product ion spectra with greater resolution and higher mass accuracies for product ions [59].

QqLIT mass spectrometer

Triple quadrupole (QqQ) is a commonly used instruments for tandem mass spectrometry; this instrument is constituted by two resolving quadrupole mass filters (Q_1 and Q_3) separated by a fragmentation cell (q_2). Triple-quadrupoles are considered to be "*beam-type*" mass spectrometers since they are usually used with a continuous ion beam source; this is both a strengthens and a weakness of these instruments; in fact when one of the mass filters is scanning, most of the ions are rejected by the instrument, resulting in poor sample use and low sensitivity [60], whereas when the mass filters are not scanning, but are transmitting a pre-selected ion, sensitivity and sample use are very high. These features, coupled with a high dynamic range, are the reasons why triple quadrupole mass spectrometers are among the most common instrument for mass spectrometric analysis of complex mixtures [61]. Moreover triple quadrupole systems are particularly interesting because they allow to perform unique and very selective analysis using particular scan modes such as *Selected Reaction Monitoring* (SRM) and *Multiple Reaction Monitoring* (MRM), *Product Ion*, *Neutral Loss* (NL) and *Precursor Ion* (PIS) scans [62]. These specific scan modes can lead to improve the detection of peptides with PTMs. For example in the *precursor ion scan* mode the Q_1 allows the transmission of all sample ions, while the Q_3 is set to monitor specific fragment ions, which are generated by fragmentation of the molecules in the collision cell (q_2). By this approach it is possible to reduce the analysis to groups of compounds, contained within a complex mixture, which fragment to produce characteristic fragment ions; a proteomics application of this scan mode is achieved by setting the Q_3 to specifically detects the fragment at m/z -79 Da specifically generated by loss of PO^{3-} from phosphorylated peptides in MS^2 analysis [63]. Fragmentation spectra obtained with this kind of instruments are of high quality since MS and MS^2 events are spatially separated. However the switch from the MS to MS^2 mode cause the reduction of the duty cycle of the instruments from 100% to about 0.1–0.01% *per* m/z, drastically reducing the sensitivity of the analysis. Additional limits for this analyzer are the low resolution and the inability to perform MS^n experiment.

As for ion-trap analyzers, they can't be classified as beam-type instruments, since they analyze pulse of ions that are introduced into the device, trapped radially by RF and then analyzed. In comparison with triple quadrupole mass spectrometers, ion trap mass spectrometers are more sensitive because of the capacity to accumulate ions. Moreover ion trap analyzers are defined as "*tandem in time*" devices that means that with single mass analyzer experiments of MS^n may be performed, providing

detailed structural information of the analytes. However ion traps are not capable of the selective scans discussed for triple quadrupole devices. As for three-dimensional ion traps (3D-IT), they are characterized by having a limited trapping capacity, mainly because of the space-charges effects (ion-ion coulombic interactions) caused by the three dimensional nature of the RF filed that entraps ions in the center of the ion trap. In addition to this limitation 3D-IT is characterized by an inherent low-mass cut-off in the resulting spectra at about 1/3 of the precursor ion m/z because the precursor ion and the lowest mass fragments must be stable simultaneously in the device during MS^n analysis.
Some of these disadvantages have been overcome by the introduction of a new class of ion trap, namely *linear ion trap* (LIT) devices. LIT optic is based on a classical quadrupole mass analyzer in which are added two lens at the end of the roads (*entrance lens* and *exit lens*) at which are applied a stopping potential DC. In the LIT ion are stored in a linear oscillatory motion along the z axis of the device, rather than in a three dimensional fashion by the RF potentials; this leads to enhance ion storage capacity. In addition the axial ion injection avoids the interference with the RF field improving the trapping efficiency. Recently radially ejection of trapped ions has been introduced by using an auxiliary voltage (*e.g.* Thermo-Finnigan LTQ). The presence of two detectors doubles the sensitivity of the analysis [46] (figure I.6).

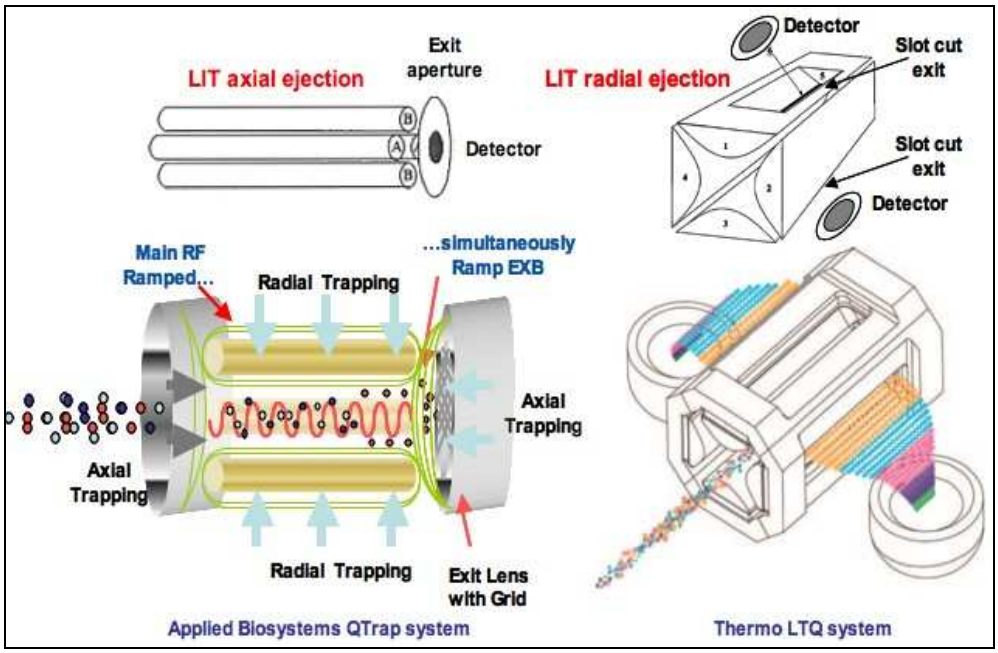


Figure I.6. *Comparison between LIT produced by Applied Biosystems and Thermo. Both the devices are characterized by the axial injection of the ions. The Applied Biosystems LIT is characterized by axial ejection and may be rapidly converted in a classical quadrupole consenting to insert it in a QqQ optic. Thermo equips their FT-ICR and Orbitrap mass spectrometers with the more sensitive radial ejection LIT. The double detection and the emptying efficiency make the Thermo LTQ a stand-alone analyzer.*

These devices usually have an ion storage capacity of about 10 folds greater than the 3D-IT devices. Recently LIT with radial ejection has been used as stand-alone mass analyzer [64]. An advantage of LIT is that trapped ions may also be ejected axially [65], by applying deformed fringing RF field to the exit lens of the LIT; this phenomena converts the motion of the ions along the plan x, y in an axial motion for that ions that are at boundary stability conditions of Mathieu diagram. Thus ions can

come across the DC barrier of the exit lens and can be selectively axially ejected changing RF potential, which brings different m/z species at the boundary conditions. However axially ejection has a low efficiency respect of radially ones (20% of ions are detected), because of the short-range effect of the fringing field. However, even if the axial ejection is less sensitive than the radial ejection, these systems make possible to switch rapidly from the LIT configuration into a classical quadrupole device.
These feature has been used from Applied Biosystems to design an hybrid mass spectrometer, commercialized with the name "*QTRAP®*", consisting of a triple quadrupole optic in which the Q_3 can operate as an axially ejection LIT [66] (figure I.7).

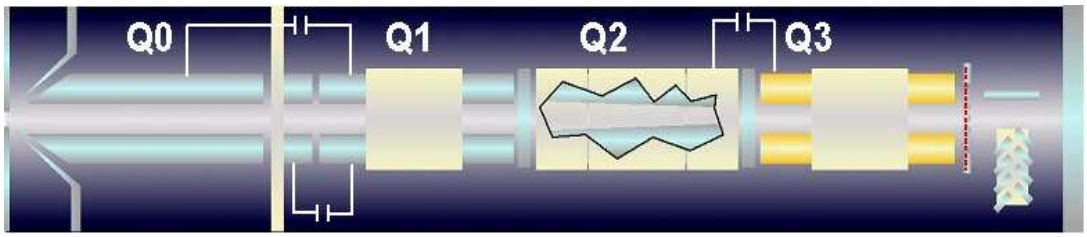


Figure I.7. *Schematic representation of a QTRAP® mass spectrometer with QqLIT optic.*

This combination combines all the vantages of a triple quadrupole with that of LIT [66]; in fact this device makes possible to perform all the selective scan modes of the triple quadrupole instruments with LIT sensitivity and high scan rate. Moreover there is no inherent low-mass cut-off in the MS/MS spectra because the fragmentation and the ion trapping steps are spatially separated [67]. Using Q_3 as a LIT it is also possible to perform MS^n analysis, taking advantage of the characteristics of higher trapping efficiency of LIT upon 3D-IT and reduced space-charge interactions that limit the number of ions that can be stored. Moreover Q_3 is operated at a very low pressure (3×10^{-5} Torr) that insure that the ions entering the LIT are collisionally cooled thereby enhancing trapping efficiency and sensitivity of MS^n. This instrument can be operated both as triple quadrupole and QqLIT on a chromatographic timescale because the switch from QqQ to QqLIT modes is accomplished by changes in electrical voltages that occur in less than 1 ms. Moreover the QTrap mass spectrometer, and in particular the 4000QTrap device, allows to perform chromatographic compatible MS^3 experiments too. The space separated fragmentation events and the high scan rate features of 4000QTrap lead to very rapid MS^3 scans (0.3-0.5 sec) allowing to carry out this kind of experiments in a LC-MS/MS method.

QqTOF mass spectrometer

Quadrupole-time of flight (QqTOF) mass spectrometers have rapidly been embraced by the scientific community as powerful and robust instruments with unique capabilities. QqTOF tandem mass spectrometer can be described as a triple quadrupole with the last quadrupole section replaced by a TOF analyser. It combines the high performance of TOF devices, in both the MS and MS/MS modes, with the widely used techniques of electrospray ionization. QqTOF mass spectrometers are characterized by high sensitivity, high mass accuracy and high mass resolution for both precursor and product ions, and also by the simplicity of operation [68-69]. This advantage does not inherently apply to the more specialized scan modes as the precursor ion, neutral loss, and multiple reaction monitoring of triple quadrupole systems, and until recently these scan methods could not be performed with any

reasonable efficiency. The popularity of the QqTOF has been significantly advanced by the rapid growth of semi-automated instrument control and data processing and by continuing improvements in the core performance characteristics of mass resolution and sensitivity. In the usual QqTOF configuration, an additional RF only quadrupole Q_0 is added to provide collisional cooling and focusing of the ions entering the instrument. So the instrument (figure I.8) consists of three quadrupoles, Q_0, Q_1 and q_2, followed by a reflecting TOF mass analyzer with orthogonal injection of ions [69].

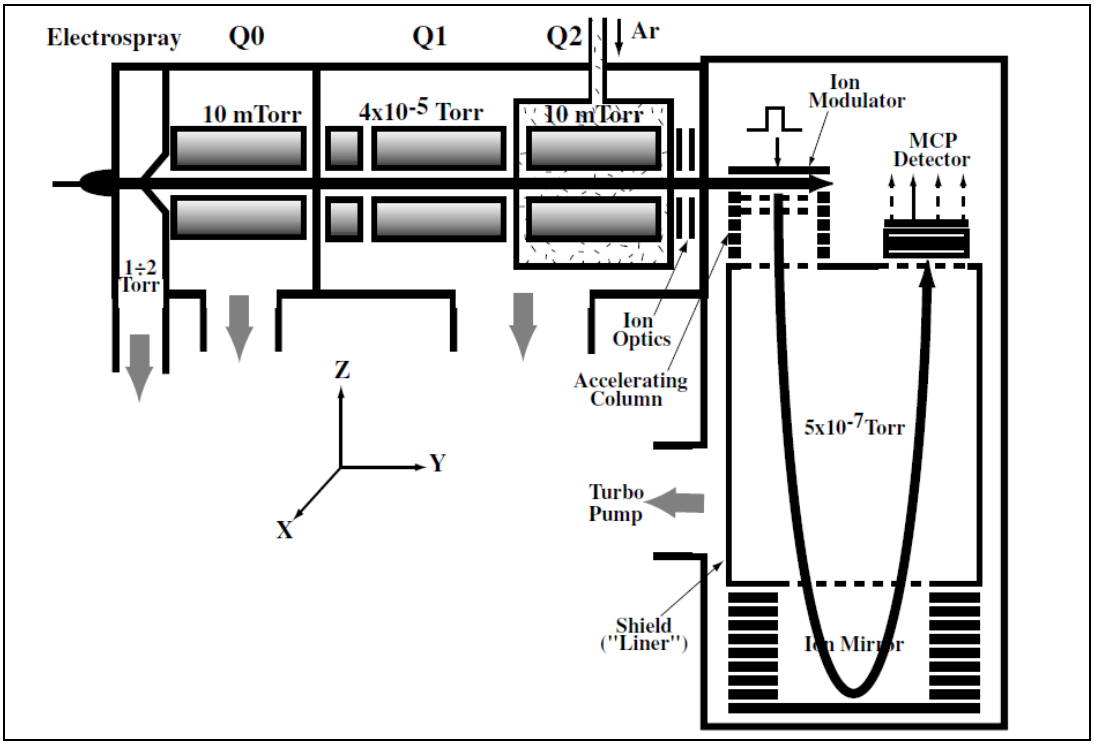


Figure I.8. *Schematic representation of a QTRAP® mass spectrometer with QqLIT optic.*

For single MS measurements, the mass filter Q_1 is operated in the RF only mode so that it serves simply as a transmission element, while the TOF analyser is used to record spectra. The resulting spectra benefit from the high resolution and mass accuracy of the TOF instruments, and also from their ability to record all ions in parallel, without scanning. Both Q_0 and q_2 are operated in the RF only mode: the RF field creates a potential that provides radial confinement of the precursor and/or fragment ions. Since the RF quadrupoles are normally operated at a pressure of several millitorr, they provide both radial and axial collisional damping of ion motion. Ions are thermalised in collisions with neutral gas molecules, reducing both the energy spread and the beam diameter and resulting in better transmission into and through both the quadrupole [70] and TOF [71] analysers.

For MS/MS, Q_1 is operated in the mass filter mode to transmit only the parent ion of interest. The ion is then accelerated to an energy of between 20 and 200 eV before it enters the collision cell q_2, where it undergoes collision induced dissociation (CID) with neutral gas molecules (usually Ar or N_2). The resulting fragment ions are collisionally cooled and focused by RF fields. This step is even more important in QqTOF instruments than it is in triple quadrupoles because the TOF analyzer is much more sensitive to the “quality” of the incoming ion beam than is Q_3 in a triple

quadrupole instrument. Both sensitivity and resolution benefit from the additional collisional focusing in the pressurized collision cell. After leaving the collision cell, ions are re-accelerated to the required energy (usually several tens of eV per unit charge), and focused by ion optics into a parallel beam that continuously enters the ion modulator of the TOF analyzer. Initially the modulator region is field-free, so ions continue to move in their original direction in the gap. A pulsed electric field is applied at a frequency of several kHz across the modulator gap, pushing ions in a direction orthogonal to their original trajectory into the accelerating column, where they acquire their final energy of several keV per charge. From the accelerating column, ions arrive in the field-free drift space, where TOF mass separation occurs. The ratio of velocities (or energies) in the two orthogonal directions is selected such that ions reach the ion mirror and then the TOF detector naturally, without requiring an additional deflection in the drift region, which could affect the mass resolution [72-73].

I.2.2 The 1st generation proteomics

Proteomics was built around the two-dimensional gel electrophoresis (2D-GE). The idea that multiple proteins can be analyzed in parallel grew from two-dimensional gel maps [25,74-75]. Initial efforts to analyse proteomes had to deal with the high complexity of samples which exceeded the capacity of analytical systems, normally obtaining limited amount of information [76]. 2D-GE was the first analytical tool successfully applied to proteomics analysis; in fact even if it was developed in the late 70s [77] it has the capability to resolve complex protein mixtures and to give protein maps from which it is possible to detect differences in protein expression patterns between different samples.
From the point of view of separations technology, 2D-GE is an orthogonal separation technique in which proteins are separated by two different physico-chemical properties: isoelectric point and molecular mass. Proteins are first separated on the basis of their (pH-dependent) net charges by isoelectric focusing (IEF) and further separated on the basis of their molecular masses by denaturing electrophoresis (SDS-Page). Both procedures are carried out in polyacrylamide gels.
Despite 2D-GE had the highest resolution power for protein separation, it still suffered from the inability to identify protein spots from 2D gels; so the foundation of the field of proteomics had to wait three technological advances of great importance: i) the introduction of large-scale nucleotide sequencing of both expressed sequence tags (ESTs) and genomic DNA; ii) the development of mass spectrometers able to ionize and mass-analyse biological molecules; iii) the development of computer algorithms able to match MS and MS/MS spectra with translations of the nucleotide sequence databases. The ability of MS to identify peptides and proteins on the base of their molecular weight [78], joined with the high resolving power of 2D-GE, was the basis for the born of proteomics.
While 2D-GE is an ideal tool for discovery-phase research, not all expressed proteins can be displayed in a single gel. Specific classes of proteins have long been known to be excluded or under-represented in 2D gels. Low-abundance proteins are not detected at all due to the limited sensitivity of commonly used dyes, or they are masked by higher abundant co-migrating proteins; excessively large and excessively small proteins, basic and acidic proteins, and membrane hydrophobic proteins present their own special challenges. Furthermore, difficult automation and reproducibility problems still remain. Some experimental improvements and technical advances, including more sensitive staining methods [79-80], large-format higher

resolving gels [81] and sample fractionation prior to 2D-GE have alleviated, but not eliminated, these drawbacks. 2D gels has historically been used to try to solve the problems of protein quantification; in fact comparative 2D gel approach has been typically used to identify proteins differential expressed (up- or down- regulated) in different samples as for example in a disease state. For this purpose a reliable analysis of quantitative changes of protein spots is important [82]; new programs for comparison of 2D-GE 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.2.3 The 2nd generation proteomics

2D-GE method still presents some limitations such as the difficulty to analyse membrane proteins because of their poor solubility their low recovery after the extraction procedures; less abundant proteins don't reach the detection limits because only low quantity of sample can be analyzed; the PTMs identification results very difficult because of the sub-stoichiometric nature of these biological modifications. For these reasons several techniques have been developed with the aim of enabling more comprehensive, less biased proteome analyses based on liquid separations coupled online with mass spectrometry, the so-called "*gel-free*" approaches (*shotgun proteomics*). For these approaches, protein samples are usually hydrolysed (most common using trypsin) to obtain more complex peptide mixtures, then subjected to liquid chromatographic separation steps prior to MS analysis. A powerful MS approach is the use of tandem mass spectrometry (MS/MS) to induce fragmentation of individual tryptic peptides after online liquid phase separation. Because of the high complexity of peptide mixtures obtained, a single-dimension chromatography does not provide sufficient peak capacity to separate all analytes; one approach to obtain high-resolution separation of peptides is using multidimensional chromatographic approaches; most common approaches use micro-capillary format liquid chromatography (μLC), in an orthogonal fashion, combining strong cation exchange (SCX) and reverse phase chromatography [33,83-85], in order to separate peptides firstly by their electrostatic charge features and then by their hydrophobicity, prior to online MS/MS analysis.
Even if progresses in MS instrumentations, as the introduction of high scan rate mass spectrometers and the realization of data depend method of acquisition, have made possible to analyse a large number of peptides in a single analysis, it is not enough to compensate the large dynamic range of protein concentration in biological samples. In order to reduce the complexity of such a mixture, a number of strategies able to select only a subpopulation of peptides for further analyses have been developed (figure I.9). In particular, if peptides carrying rare amino acids like cysteine, methionine or tryptophan are selected, a given protein may only contribute a single or a few peptides to the selected fraction; in most cases these remaining peptides still allow the unambiguous identification of the proteins they originate from [86]. These analytical strategies allow the identification of less abundant proteins, as for as the enrichment of peptides carrying various PTMs.

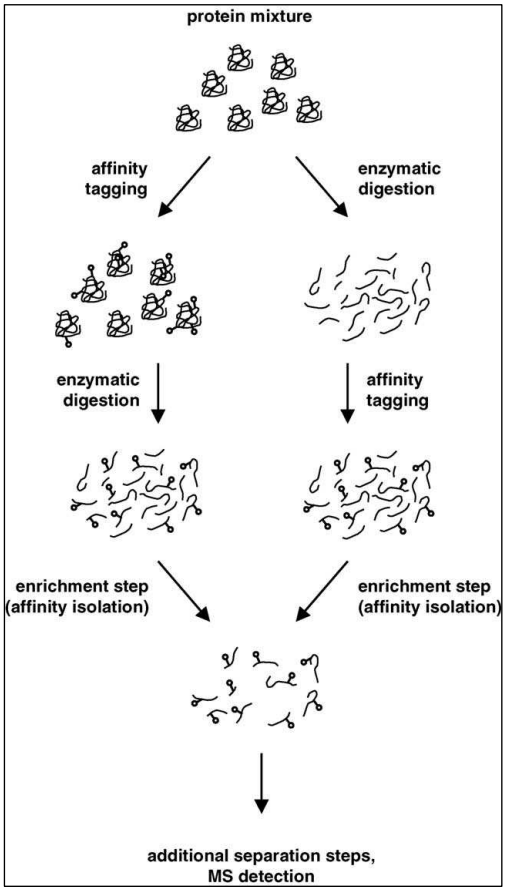


Figure I.9. *The use of chemical tagging strategies for sample fractionation. A protein mixture is either first labeled with an affinity tag and then digested (left) or first digested and then labeled (right). In both cases, labeled peptides are subsequently enriched by an affinity chromatography step, so that ideally only the tagged peptides remain.*

The first innovative methodologies called ICAT (Isotope Coded Affinity Tag) was introduced by Aebersold's group in 1999 [87]. They based their strategy on the enrichment and the analysis of the only fractions of peptides containing cysteine residues. It was chosen this amino acid because of the evidence that it is present in the 92% of *Saccharomyces cerevisiae* proteins and by their hydrolysis with trypsin only the 10% of the resulting peptides contains it [88]. The global strategy consists in the labelling of cysteine residues with a thiol reactive biotin and the sequent affinity enrichment for cysteine containing peptides by streptavidin affinity chromatography prior to 2D-LC-MS/MS analysis. On these basis many others labelling strategies had been developed aimed to enhance both the global protein identification and more specific PTMs.

I.2.4 The 3rd generation proteomics

Third generation proteomics represents a progress in simplification processes of proteomics analysis. In fact it aims not only to eliminate the use of 2D-GE as a separation step, but also to avoid the enrichment step, required to reduce peptide mixture complexity, prior to LC-MS/MS analysis. The guidelines of this methodology are: i) selective labelling of residues of interest; ii) employment of particular derivatizing reagents that supply characteristic ions of particular stability in fragmentation; iii) selective analysis by tandem mass spectrometry. This analytical technique has been called RIGhT (Reporter Ion Generating Tag) strategy [89]. It is included in the class of analytical methodologies that require the use of hybrid instruments for MS analysis, by which it is possible to carry out analysis in *precursor ion scan* mode. PIS allows the selective analysis of the only precursor ions that generate the specific fragment of interest; this represents a great advantage in terms of analysis sensitivity. Interesting examples of RIGhT methodology are its use for selective identification of pSer/pThr containing peptides [89], for identification of 3-nitrotyrosine residues [90] and for selective identification of carbonylated lysine residues [91].

I.2.5 Quantitative proteomics

The field of proteomics has moved beyond simple protein identification and is now addressed to accurately and reliably quantitate the differences and the relative variation in protein abundance in an organism, cell, or tissue, at a given time or under a particular condition [92-94]. Traditionally, quantitative analysis has been exploited

using a so-called differential display 2D-GE approach, whereby differences in protein abundances were determined by comparing stained protein spot volumes, followed by identification of proteins by mass spectrometry. This approach suffers from the disadvantages noted above for the 2D-GE and, in addition, spot intensities can be affected by limited precision because of sample handling and sample processing variability.

More recently quantitative proteomics has moved beyond the 2D-GE approach and can be separated into two major approaches: the use of stable isotope labelling and label-free techniques.

As for stable isotope labelling, the most common strategies involves labelling of peptides with iTRAQ (isobaric tag for relative and absolute quantification) [95] or TMT (tandem mass tag) [96], labelling of proteins with ICAT (isotope-coded affinity tag) [86], using the serine protease catalysed attack of ^{18}O water in the digestion step [97-98] or using SILAC strategy for metabolically labelling proteins by incorporation of isotope labels in cell cultures [99]. These strategies are considered to be more accurate in quantitation, but they require expensive molecules, specific software and expertise to analyse data. Moreover the number of samples that can be contemporary analysed is limited by the number of available labels and not all strategies can be applied to all types of samples. As for labelling strategies, targeted towards specific amino acid residues (cysteine, histidine and methionine) or N- or C-terminal, they permit both the enrichment of subfractions via affinity clean-up and the quantitation leading contemporary to the identification of an ever increasing number of proteins and to their quantitative analysis [100]. A typical approach based on the isotope stable affinity tagging and mass spectrometric analysis is isotope coded affinity tag (ICAT) methodology [87]. The reagent consists of: i) a cysteine reactive group (an iodoacetamide moiety); ii) a linker that contains either a group of heavy or light isotopes; iii) a biotin affinity tag that allows the enrichment of labelled peptides via streptavidin affinity chromatography (figure I.10). The light and heavy reagents are chemically identical but differ in masses. ICAT reagent is commercialized by Applied Biosystems in two different isotopic forms with a different distribution of $^{12}C/^{13}C$ atoms that allows the relative quantitative estimation between two samples. Labelled proteins are firstly combined and then digested with proteolytic enzyme, usually trypsin. Labelled peptides (namely cysteine-containing peptides) are isolated using avidin affinity chromatography and subsequently identified and quantified by MS. The ratio of protein between the two samples can then be measured by the peak ratio between the “light sample” and the “heavy sample” [101].

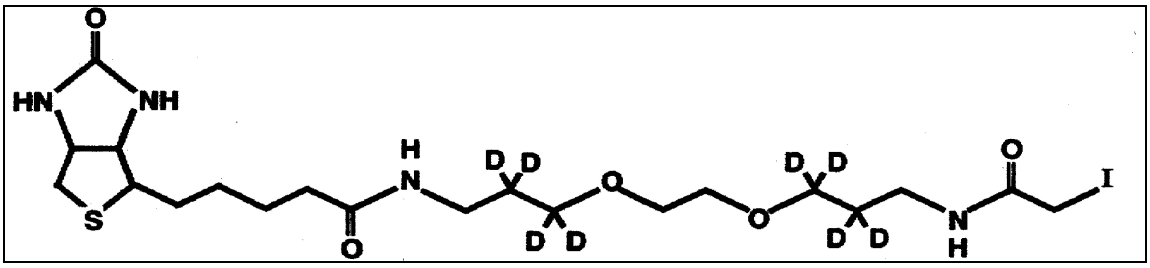


Figure I.10. *Structure of a common ICAT reagent.*

A more recent methodology for quantitative analysis by means of mass spectrometry, concerns the use of iTRAQ reagents. iTRAQ technology has been developed and commercialized by Applied Biosystems and allows to quantitate proteins by MS^2

analysis rather than by MS analysis [95]. The structure of iTRAQ reagents is shown in figure I.11; the complete molecule consists of: i) a reporter group (N-methylpiperazine moiety), that retain the charge and provides a stable signal during MS/MS analysis; ii) a mass balance group (carbonyl moiety), lost as neutral molecule; iii) a peptide-reactive group (NHS ester).

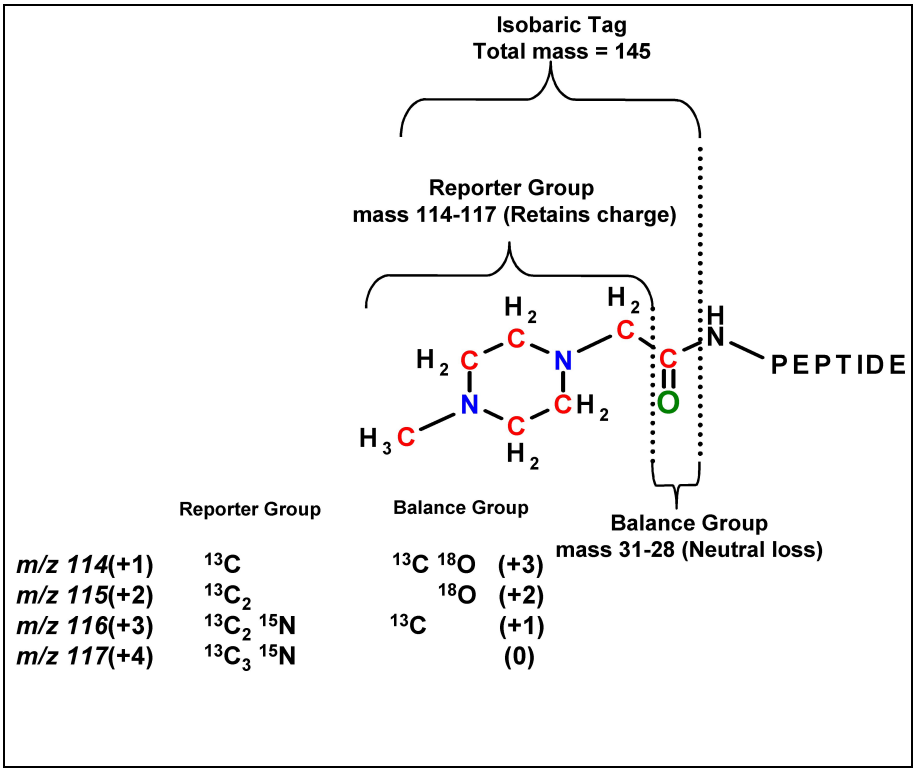


Figure I.11. *Structure and isotopic composition of iTRAQ reagents.*

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 (figure I.11). 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 reagents, reactive toward primary amino groups, are differentially isotopically labelled such that all derivatized peptides are isobaric and chromatographically indistinguishable, therefore also identical in single MS mode. Upon fragmentation, the reporter groups are detached, creating signals in the low mass range of the MS/MS spectrum; thus low mass reporter ion signals allow quantitation, while peptide fragment ion signals allow protein identification [95].

I.3 Post-translational modifications

Post-translational modifications (PTMs) are chemical or enzymatic modifications of a protein after its synthesis. Post-translational modification of proteins is part of what makes proteomics so much challenging. Amino acids can be modified by literally hundreds of modifications that change their molecular weights, the fundamental physical property measured by mass spectrometry. Clearly, post-translational modification of proteins is an event with dramatic effects on the complexity of the proteome. Post-translational modification of amino acids extends the range of

functions of proteins by attaching them other chemical functional groups, by changing chemical nature of amino acids or by making structural changes [102].
Nowadays the attention towards PTMs is notably increased due to the fact that they play a critical role in cellular functioning and they vary in response of environmental stimuli, widespread or development of diseases. Phosphorylation, acetylation, nitration, are only some examples of post-translational modification involved in important cellular processes [103-105]. For these reasons they are object of study in proteomics laboratories. New methodologies in qualitative proteomics, such as "shotgun proteomics", and in quantitative proteomics, have the capability to be applied in rational strain improvement projects and could be considered the starting point for industrial bioprocesses improvement and for the design of novel bioprocesses as for as for the design of new producer strains [106-107]. PTMs play an important role for rational strain improvement since they are the main way in which cells regulate their enzymatic and metabolic activity. Protein phosphorylation is a typical example of a PTM involved in the modulation of enzymatic activity as for as in signal transduction pathways. PTMs linked with oxidative stress phenomena, are also important in modulation and fine optimization of productive processes that involve microorganism; this PTMs are mainly due to the accumulation of ROS (Reactive Oxygen Species) and RNS (Reactive Nitrogen Species) in the cytosol and in the broth, with the progress of fermentative production processes [108-110]. In a perspective such as that of systems biotechnology, the global study of the mechanisms by which the presence of these oxidizing species and the phenomena of enzymatic activity regulation bearing on efficiency of producer strains, plays a key role in the design of new producer strains through targeted strategies of cellular engineering.

Proteomics for studying PTMs

Proteomics is one of the disciplines of choice to study these mechanisms, and a number of strategies of analysis for the study of PTMs are already of common use. Unfortunately, the efforts made in this direction have produced only lists of modified proteins, struggling to demonstrate the correlation to observed phenotypes. There is so much to do in this area, particularly to develop methods of quantitative analysis to be exploited for the development of new biotechnological processes.
Most of the common strategies for the analysis of PTMs exploit the potential of 2D-GE [111-112] coupled to Western blot analysis [113]. However, 2D-GE has well known drawbacks that limit the analysis of post-translationally modified proteins. A major limitation concerns the ability to pinpoint the modification site that is fundamental in order to design a strategy of targeted engineering of proteins. Moreover it is not always possible to have quantitative information about selected modified residues that are fundamental to compare different growth conditions. For these reasons it is preferable to use liquid chromatography-electrospray ionization methods (ESI-LC-MS/MS) [114], although PTMs analysis from gels is sometimes the only option. There are some problems related to mass spectrometric analysis of PTMs; first of all the modified peptide must be detected. In fact, even if it is almost always possible to identify a protein from MS/MS spectra of only a few of its peptides, locating the exact site of a modification requires detection, and in most cases MS/MS sequencing, of the specific modified peptide. This can be difficult since not all peptides are equally detected; moreover not all species are equally abundant in a sample and in particular

post-translationally modified ones are usually less abundant than the non modified species.

I.3.1 Protein phosphorylation

Phosphorylation is the addition of a phosphate group to specific amino acid residues. Proteins and peptides phosphorylation is one of the most important and abundant PTMs and a biochemical process of absolute biological relevance; in fact kinases and phosphatases, the enzyme responsible for protein phosphorylation and dephosphorylation, account for 2-4% of eukaryotic proteomes [115-116]; phosphorylation cycles are very dynamic and rapid in order to satisfy necessity of the cell for highly dynamic regulation processes. Phosphorylation acts as a molecular switch, in a reversible manner, to turn “on” or “off” a protein activity or a cellular pathway, regulating enzymatic activities or substrate specificities, folding and function of proteins, protein localization, complexes formation and degradation. Cellular processes ranging from cell cycle progression, differentiation, development, hormones response, signal transduction, metabolic maintenance and adaptation, are all regulated by protein phosphorylation. Recent estimations of proteins phosphorylation, suggest that almost one-third of eukaryotic proteins are phosphorylated [117-118].
One of the major problems concerning the analysis of this PTM is the rate of phosphorylation; in fact phosphoproteins abundance is very low and can be quantified as 1-2% of the total protein amount in a cell. So protein phosphorylation can be classified as a sub-stoichiometric modification and usually, while some residues are always phosphorylated, others may only be transiently phosphorylated. As for modification sites, we have to distinguish different kinds of protein phosphorylation, on the basis of the amino acid residues involved; we can classify four different types of phosphorylation: i) O-phosphorylation is the most common class and in this case phosphate groups are mostly attached to serine, threonine and tyrosine residues, but also to unusual amino acids such as hydroxyl-proline; ii) N-phosphorylation occurs mostly on histidine and lysine (N-phosphates); iii) S-phosphorylation occurs on cysteine (S-phosphates); iv) acyl-phosphorylation occurs on aspartic and glutamic acid residues (acyl-phosphates). Except for O-phosphates others classes of phosphorylation are less common. In many cases, the effects of phosphorylation are combinatorial and multiple sites could be phosphorylated contemporary.
For all of these reasons the definition of sites, abundances, and roles of each of these modifications in a biological sample is a challenging task for proteomics, mainly because they are not accessible to a single analytic method altogether at the same time.

Phosphoproteome analysis

To analyse protein phosphorylation the first step is the correct sample preparation, in order to avoid that protein phosphatases will become active and phosphate groups will be lost [119]. Several methods have been developed for detecting phosphoproteins. 2D-GE has been used for this purpose [120-121]. After electrophoretic separation proteins can be classically visualized using standard protein stains, as Coomassie or silver staining; with this strategy the modified proteins could not be univocally identified because they could appear at different pI according to the nature of the

phosphorylation (addition or subtraction of negative charges along the molecule). This is especially evident for multiply phosphorylated proteins.
Phosphoproteins separated by 2D-GE can also be visualized using phosphospecific stains such as Pro-Q Diamond [122-123], a fluorescent dye that allows to detect 1-20 ng of phosphoprotein, depending on the phosphorylation state of the protein. For individual phosphoproteins, the strength of the signal correlates with the number of phosphate groups. Using a specific staining procedure, this method can be used to specifically detect Ser/Thr- phosphorylations.
Phosphoproteins on 2D gels can also be visualized by Western blotting using special antibodies against phosphoserine, phosphothreonine or phosphotyrosine residues [124]. To measure changes in protein phosphorylation state, two distinct proteome samples can be run and immunoblotted and the intensity of spots can be compared. All these methods have a great drawback in the inability to provide information on modification sites. Protein spots of interest must be excised from gels and the proteins analysed by MS, but in a high number of cases this analysis can be difficult and less useful due to co-migrating proteins. Another classical approach to study phosphoproteins is the radioactively labelling using ^{32}P or ^{33}P [123,125] and their detection by autoradiography; this method is the most sensitive available, but relies with handling radioactive compounds and therefore is not the right choice for a high throughput strategy. Moreover radioactive labelling is not compatible with classical proteomics procedures first of all MS analysis.
Radioactive labelling is possible both *in vivo* and *in vitro*. In *in vivo* studies, cells are incubated with ^{32}P; in some cases, the presence of endogenous ATP can interfere with the incorporation of the label resulting in inefficient radioactive labelling [126]. In *in vitro* studies, proteins are incubated with specific kinases in the presence of [γ-^{32}P]-ATP, and under appropriate conditions ^{32}P is incorporated at the phosphorylated amino acid residue. The major drawback of this method is the possibility that kinases may phosphorylate proteins aspecifically, due to their unnatural high concentration [127]. The quantitative analysis in this case is performed by comparing differences in relative abundance of ^{32}P-labeled phosphoproteins, visualized by autoradiography, by measuring relative intensities of the spots on the gel.
In order to increase sensitivity, selectivity, speed and to eliminate the need for radioactive labelling and antibodies, several strategies have been developed, the majority of which are mass spectrometry based, for the analysis of phosphorylation. MS based methods are mostly based on the analysis of phosphopeptides obtained by proteolytic processing, rather than on the analysis of phosphoproteins. However, phosphopeptides are difficult to analyze by MS, especially in the presence of non modified peptides, mainly because of their low ionization efficiency. Thus, efficient enrichment of the phosphorylated species prior to MS analysis is a critical step to increase sensitivity of the method and to achieve a more efficient characterization. Different enrichment strategies have been developed, and still continue to evolve. The most frequently used enrichment strategies are affinity-based, such as immobilized metal affinity chromatography (IMAC) and metal oxide affinity chromatography (MOAC), or strong cation exchange (SCX) chromatography.
IMAC [128] is the classical phosphopeptide enrichment technique and is based on the principle that positively charged metal ions, such as Fe(III) [129-130] or Ga(III) [131], chelated onto a solid phase resin, interact with negatively charged phosphate groups at acidic pH; subsequently phosphopeptides are eluted at basic pH. Some protocols make use of a derivatization step to convert carboxylic acids to their corresponding methyl esters prior to IMAC to prevent non specific binding [132].

Recently MOAC protocols have been developed taking advantage of titania or zirconia as metal oxide chromatography modifiers [133]. Actually titanium dioxide (TiO_2) remains the most commonly used modifier. In contrast to IMAC, which can have greater affinity for multiply phosphorylated peptides, TiO_2–MOAC seems to be more specific towards singly phosphorylated peptides [134].
By SCX chromatography, phosphopeptides are enriched in the earlier-eluted fractions since they have lower affinity for the negatively charged resin than corresponding non-phosphorylated peptides of the same sequence.
Recent efforts have aimed to completely automate the multistep, tedious task of phosphopeptides enrichment by conducting IMAC [135], TiO_2 [136-137], or SCX [138] online with MS.
A number of alternative enrichment strategies for phosphopeptides have recently been proposed. An example is the BEMA (β-elimination/Michael addition) strategy, that takes advantage of the ease of β-elimination of pSer and pThr residues at basic pH, and the ability to subject the resulting products to Michael type addiction with a desired tag [139]; the major disadvantages of this techniques are that pTyr residues cannot be addressed and the possibility to have contamination from peptides with other modifications that also undergo β-elimination, such as O-glycosylated peptides.
In 2001, Zhou and co-workers introduced a strategy called phosphoramidate chemistry (PAC), which can be applied to all phosphorylated residues [140]. Phosphopeptides are derivatized by a sulfhydryl group and linked to iodoacetyl groups immobilized on a resin through phosphoramidate chemistry. The phosphate groups are reconstituted by acid hydrolysis (using TFA) of the phosphoramidate bonds in order to allow the identification of the phosphorylation sites by MS. However, this strategy is time consuming and sample may be lost at each step.

Mass spectrometry for the analysis of phosphorylation

In phosphorylation studies it is very important to be able to specifically identify the amino acid residue or the amino acid residues that are phosphorylated upon a certain stimulus. Using mass spectrometric analysis it would be possible to address the average number of phosphorylation sites just by looking at the increment of the protein mass compared with the unmodified one, as a phosphate group gives a mass increment of 80 Da. However, the exact location and nature of the modifying group cannot be achieved just by mass spectrometric measurement of the intact protein. Nowadays most phosphorylation sites are identified by analysing the list of observed masses of peptides obtained from tryptic digestion of protein of interest and comparing them with the list of expected peptide masses, in order to find mass increases of 80 Da. Although this method is relatively straightforward, it could miss some phosphorylated species, especially because of their low ionization efficiency (caused by the increased acidity of phosphorylated peptides) and because of the competition for ionisation of peptides in complex mixtures, which results in suppression of signals for some peptides. A useful solution could be removing phosphate group in order to enhance the relative intensity of phosphorylated species [142]. Any ambiguities can be resolved by sequencing the peptide using tandem MS [140-141]; ideally, an MS/MS spectrum of a phosphorylated peptide generates a numbers of fragment ions that differ in mass by a single amino acid, that can be used to determine the peptide sequence and position of the phosphoryl residue(s).
In order to solve the problem of competition for ionization, a powerful approach makes use of LC-ESI-MS/MS strategies [143-144]. In fact this tool has established the

basis for several novel methodologies for the identification of phosphopeptides. Whit this trick the ion suppression phenomena are reduced by separation of peptides prior to MS analysis. Moreover this method has the advantages to concentrate samples and to remove salts that could interfere with ionization. As for MS/MS analysis of phosphopeptides, the latest have characteristic fragmentation patterns that can be exploited for the selective analysis of these species. In fact, in the case of serine and threonine phosphorylated peptides, it is possible to find specific fragment ions, with a ΔM of -98 Da from parent ions, due to the loss of H_3PO_4 as a neutral molecule [145-147]. This characteristic fragmentation behaviour can be used in MS/MS experiments, by using QqQ mass spectrometers, to selectively analyse phosphorylated peptides, using the neutral loss scan mode typical of these instruments.

I.3.2 PTMs induced by oxidative stress

In living cells, following oxidative metabolism, oxygen molecules are normally reduced to produce water molecules, a reaction catalyzed by cytochrome oxidase as the terminal part of the electron transport chain; a minimal part of oxygen (1-2%) [148-149] can be reduced partially with the formation of partially reduced by-products, such as superoxide anion (O_2^-•) and hydroxyl radical (•OH), collectively named ROS, *reactive oxygen species*. These molecules are characterized by an unpaired electron (in fact they are also called free radicals) and are extremely reactive toward the surrounding molecules in an indiscriminate fashion; for this reason these molecules are short-lived and can be harmful if they are not controlled. The major source of ROS is the electron transport chain in mitochondria [150-151]. In physiological conditions ROS have an essential role in cell signaling [152-154] and their activity is normally tightly regulated by a thick network of antioxidants, both enzymes and small organic molecules. A significant increase in the levels of intracellular oxidizing species causes an imbalance between antioxidant systems and ROS that leads to the onset of oxidative stress conditions; this state is frequently connected to pathological conditions [155-158]. However, in most disorders, oxidative/nitrosative stress is a consequence and not a cause of the primary disease process. When superoxide anion increases its cellular concentration, it is rapidly converted to hydrogen peroxide (H_2O_2) and oxygen (O_2) by superoxide dismutase (SOD), that is the main regulator of cellular concentration of O_2^-• [159]. H_2O_2 is then removed by catalase and peroxidase that convert it in water. If H_2O_2 accumulates in the cell, it can react rapidly with transition metals giving rise to the hydroxyl radical (•OH), a very reactive and toxic specie [160]. In addition to producing ROS, the mitochondrial respiratory chain can produce nitric oxide (NO•), which has itself an unpaired electron, and other nitric oxide by-products called RNS, *reactive nitrogen species*. RNS play many physiological roles, mainly as signal molecules [161], but if they are present in high concentration may induce strong modifications in many biological macro-molecules [162]. Nitric oxide is a ubiquitous intercellular messenger that modulates blood flux and neural activity and it is also important for nonspecific host defense, helping to kill tumours and intracellular pathogens. Nitric oxide by itself is not a toxic compound; the toxicity of nitric oxide results from its reaction with ROS to produce a class of derivatives extremely reactive towards proteins, lipids, and DNA [163]; for instance when nitric oxide reacts with superoxide anion it is produced peroxynitrite ($ONOO^-$) a powerful and toxic oxidant [164].
Obviously both enzymatic and non-enzymatic systems protect cells and tissues against the toxic effects of ROS/RNS produced by metabolism. However, in particular

cases, such as inflammatory states or during fermentative processes [165], the concentration of such species may exceed the cell antioxidant skills making relevant the toxic effects of these molecules.
Post-translational modifications of proteins by ROS and RNS are associated to a large number of pathologies and to cellular aging [156,157]. Some of these PTMs affect the structural and functional integrity of specific proteins. Most of such oxidants will react in a non specific fashion with several amino acid residues, yielding to multiple products. However, only a few of these reactions will lead to stable and specific products, characteristic for a selected oxidant, which can be used as a signature for the reaction of one specific oxidant with a protein *in vivo*.

I.3.2.1 Protein nitration

The presence of peroxynitrite in cells can cause nitration of amino acid residues. Targets of nitration include nucleic acids, sugars, lipids, alcohols, small organic molecules and aromatic amino acids [166]. Nitration of proteins modulates catalytic activity, cell signalling and cytoskeletal organization. Sites of nitration in proteins are phenylalanine, tryptophan and, in particular, tyrosine residues; in fact the phenolic ring of tyrosine is particularly active towards this kind of reaction. Nitration of tyrosines is a stable PTM that reflects the extent of the production of oxidants during both physiological and non-physiological conditions. For this reason nitrated tyrosines have been recognized as the prevalent and functionally significant PTM that serves as an indicator of RNS mediated oxidative reactions [167].
There are two different catalytic mechanisms proposed for *in vivo* nitration (figure I.12), both involving the formation of tyrosyl radical [168], generated by the presence of some ROS, such as CO_3^{-}• and OH•, mainly formed by peroxynitrite [168]. Tyrosyl radical reacts with nitrogen dioxide to generate NO_2Tyr. Tyrosyl radical can also reacts with another tyrosyl radical to yield the 3-3'-dityrosine specie, which can generate intramolecular cross-links or protein multimerization. These pathways might be less important *in vivo* compared with *in vitro* observation because of the presence of tyrosyl radical reductants, steric hindrance and tyrosyl radical low accessibility *in vivo*.
Another possible way for nitration of tyrosines *in vivo* is the reaction of tyrosyl radical with nitric oxide (NO•), resulting in the formation of 3-nitrosotyrosine. Nitrosotyrosine is apparently not stable and it is oxidized by two electrons to form 3-nitrotyrosine (figure I.12) [169].
Addition of an NO_2 group to the orto position of tyrosine residues decreases the pKa of the tyrosine hydroxyl group of about three units [170]; this introduces a net negative charge to approximately half of nitrated tyrosine residues at physiological pH. This new anionic amino acid residue can induce changes in protein conformation and biological activity and it can generate antigen epitopes; moreover the altered pKa of hydroxyl groups can cause inhibition of tyrosine phosphorylation by protein kinases. Nitration of tyrosine seems to have certain selectivity; in fact not all tyrosine residues are nitrated; tyrosine nitration usually occurs near acidic residues, in loop regions [171-172] or in areas free from steric hindrances. Tyrosine nitration sites are localised within specific functional domains of nitrated proteins [171].

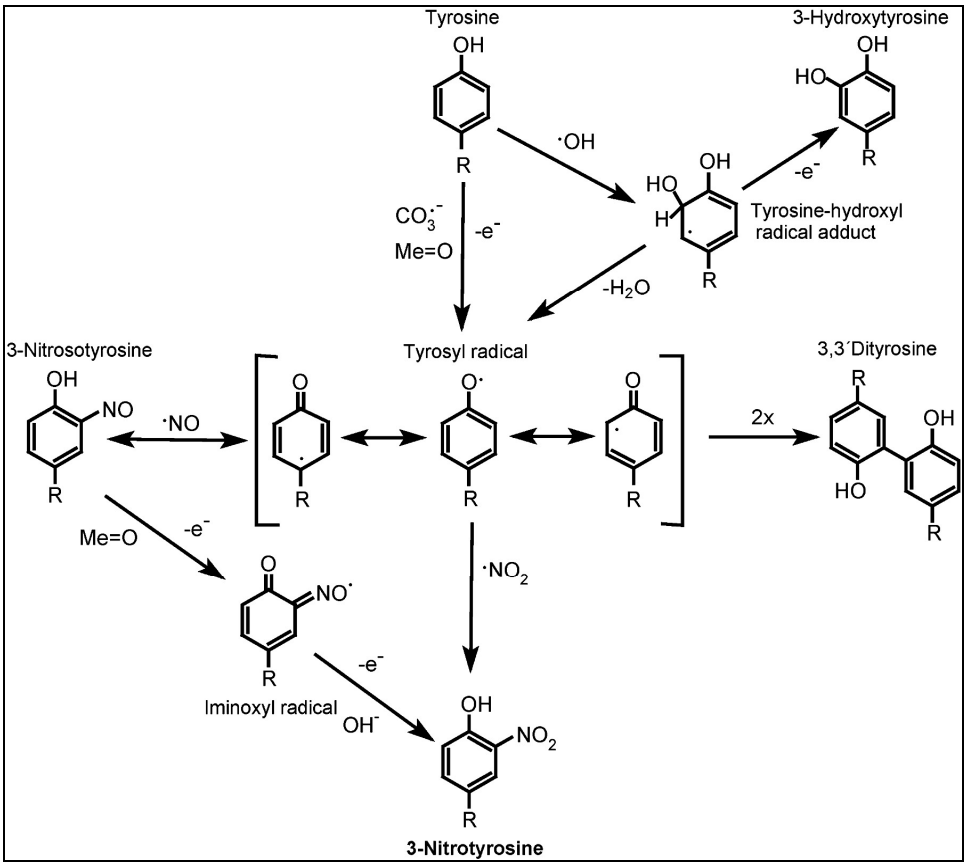


Figure I.12. *Main pathways of in vivo generation of 3-nitrotyrosine residues.*

Nitration of tyrosine is one of the sub-stoichiometric PTMs related to ROS/RNS that is stable *in vivo* and it is directly linked to the instauration of oxidative stress conditions caused by the iper-production of nitric oxide [173]. For these reasons NO_2Tyr is currently considered the best biomarker for the identification and quantification of oxidative stress events related to RNS [174].

Nitroproteome analysis

Most of the studies on protein nitration published so far are focused on the detection and quantification of nitrotyrosine by using 2D-GE and Western blot [175-176], immune-histochemical or ELISA techniques [177-178], or chromatographic methods using either high performance liquid chromatography (HPLC) separation [179-181] combined with electrochemical detection or GC-MS techniques. MS can also be used to analyse overall levels of 3-nitrotyrosine in proteins; in fact the formation of NO_2Tyr residues cause an increment of 45 Da in the molecular mass of the modified peptide respect to the unmodified one. MALDI mass spectrometry has been extensively used for the analysis of nitrotyrosines; however, nitrotyrosine containing peptides are sensitive to laser light and undergo a decomposition process yielding, besides the expected mass increase of 45 Da, other major peaks corresponding to nitrosotyrosine ((Tyr-NO) [M + H^+ + 29]), aminotyrosine ((Tyr-NH_2) [M + H^+ + 15]), and nitrenetyrosine ((Tyr-N) [M + H^+ +13]) [182-183]. While on the one hand this phenomenon constitutes a unique fingerprint, it also impairs the sensitivity of the analysis, since it spreads the signal of nitrated peptide over at least three major species, hence reducing the signal intensity of nitrated specie by as much as 60-70%, which may result in the failure to

observe nitrated peptides in complex peptide mixtures [184-186]. This drawback is aggravated by the fact that tyrosine nitration is a low-abundance PTM with an occurrence of 1 in 10000 tyrosines *in vivo* [168]. It has been showed by Petersson and co-workers [183] that using ESI mass spectrometry this phenomenon is absent. Moreover it is possible to use ESI-MS/MS in *precursor ion scan* mode to detect nitrated peptides in complex mixtures using a characteristic fragment ion of nitrated peptides such as the immonium ion of nitrotyrosine [183].

I.3.2.2 Protein carbonylation

ROS generally reacts with proteins via metal-catalysed reactions (MCO) [187-188] causing cleavage of polypeptide backbone [189], intermolecular cross-linking and modification of the amino acid side chains, leading to the loss of protein function and to structural alteration [190]. MCO generally involves generation of H_2O_2 and reduction of Fe(III) or Cu(II) by a suitable electron donor, like NADH, NADPH, ascorbate, and so on; Fe(II) and Cu(I) ions bind to specific metal binding sites on proteins and react with H_2O_2 to generate OH•, an highly reactive free radical specie [191-192].
Among the various oxidative lesions caused by this radical, protein carbonylation is an irreversible and irreparable chemical post-translational modification that cannot be reversed by the enzymatic repair machinery of cells. Carbonyl derivatives are essentially formed by direct MCO attacks on the amino acid side chains of arginine (R), lysine (K), threonine (T) and proline (P) residues [193-194]. Carbonyl groups can also be introduced into proteins via secondary oxidation reactions; in fact lipid peroxidation products [195] and glycation end products [196] can react with amino acid residues via Schiff-base formation and further oxidation.
Lysine, arginine and proline residues are readily and directly oxidized by metal ion-catalyzed oxidation (MCO) systems [197-198]; oxidation of the side chains of lysine, arginine, proline, and threonine residues has been shown to yield carbonyl derivatives; oxidation products of arginine and proline residues (glutamic semialdehyde) and lysine residues (α-aminoadipic semialdehyde) (figure I.14), are very reactive groups; as for histidine residues, they are converted to 2-oxo-histidine [194]. It has been demonstrated that oxidation of proline residues can lead to peptide bond cleavage [199]. In addition it was demonstrated that the oxidation of glutamic acid and aspartic acid residues can also lead to peptide bond cleavage in which the N-terminal amino acid of the C-terminal fragment will exist as the N-pyruvyl derivative [200]. Carbonyls containing proteins can also be generated by reaction of ε-amino groups of lysine residues with reducing sugars or their oxidation products (glycation-glycoxidation reactions) [201-203] and also by Michael type addition reactions of lysine, cysteine, or histidine residues with α,β-unsaturated aldehydes formed during the peroxidation of poly-unsaturated fatty acids [204].
Approximately 10% of the proteome is more prone to carbonylation during ageing, starvation or disease [205-208]. There are several possible explanations for this specificity, as for examples the presence of a transition metal in the protein or the localisation of these proteins close to ROS generating sites; however the exact molecular basis for this apparent specificity still remain unclear [209].
Recent studies have demonstrated that only few sites are selectively carbonylated among all possible carbonylatable sites and carbonylation sites are mainly located at the protein surface [210-214]. The formation of carbonylated species in proteins seems to be a common phenomenon during oxidation due to aging or inflammatory processes, and the quantification of carbonylated species can be used to measure the extent of

oxidative damage. In fact usually carbonyls are more difficult to induce than methionine or cysteine oxidation, and might therefore indicate a more severe oxidative stress [215]. This has prompted the development of a number of analytical strategies to determine the protein carbonyl content in biological samples; a common method exploits the derivatization of carbonyl groups with 2,4-dinitrophenylhydrazine or tritiated borohydride followed by quantification with UV spectroscopy or radiography, respectively [216-218]; another method to quantify carbonyl content is by derivatizing with a fluorophore such as fluorescamine; the resulting product is fluorescent and can be revealed by spectroscopy [219].

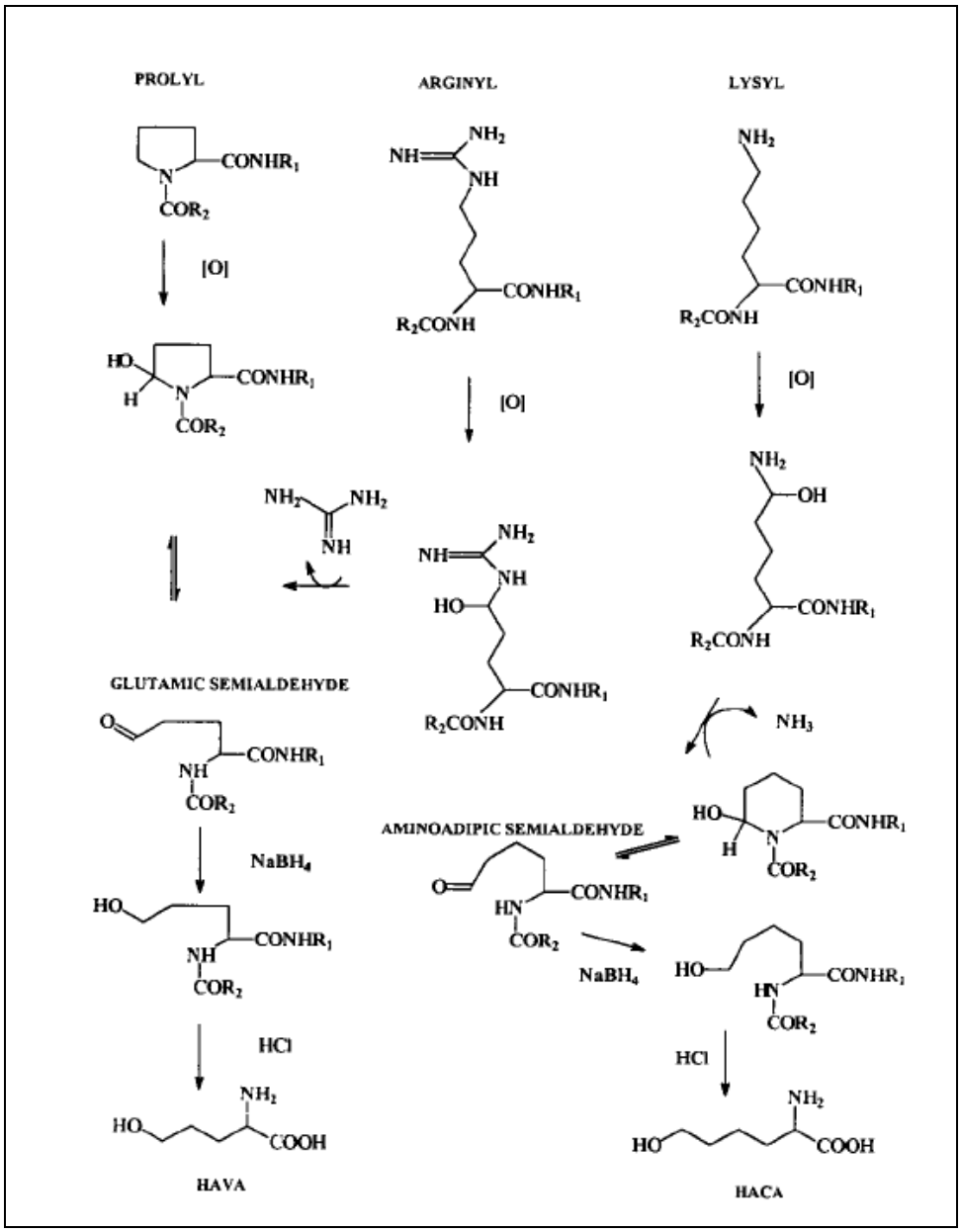


Figure I.14. *Common amino acid side chains involved in oxidation and their oxidation products.*

A number of methods are based on immunological detection of derivatized carbonyl groups; an example is the use of 2,4-dinitrophenol antibodies to detect and quantify 2,4-dinitrophenylhydrazyne derivatized proteins [220]. For identification of carbonylated proteins all these methods use mass spectrometry as final tool; although these

methods are able to quantify and identify carbonylated proteins they are not able to pinpoint sites of carbonylation; however identification of oxidation sites provides a tool for deeper understanding what oxidative stress causes in living cells.

I.4 References

1. Zhou, M.; Veenstra,T.D. *Proteomics*. **2007**, *7*, 2688-2697.
2. Josic, D.; Clifton, J.G. *Proteomics*. **2007**, *7*, 3010-3029.
3. Le Naour, F.; André, M.; Greco, C.; Billard, M. *Mol Cell Proteomics*. **2006**, *5*, 845-857.
4. Mei, J.; Kolbin, D.; Kao, H.T.; Porton, B. *Schizophrenia Res*. **2006**, *84*, 204-213.
5. Cox, J.; Mann, M. *Cell*. **2007**, *130*, 395-398.
6. Session “Biotechnology perspectives”, HUPO First world congress. *Mol Cell Proteomics*. **2002**, *1*, 709-710 and 721-725.
7. Session: “Proteomics in biotechnology”, HUPO 4th Annual world congress. *Mol Cell Proteomics*. **2005**, *4*, S285-S289.
8. Incamps, A.; Hély-Joly, F.; Chagvardieff, P.; Rambourg, J.C. *Biotechnol Bioeng*. **2005**, *91*, 447-459.
9. Josic, D.; Brown, M.K.; Huang, F.; Lim,Y.P. *Proteomics*. **2006**, *6*, 2874-2885.
10. Gupta, P.; Lee, K.H. *Trends Biotechnol*. **2007**, *25*, 324-330.
11. Kalinowski, J.; Bathe, B.; Bartels, D.; Bischoff, N.; Bott, M.; Burkovski, A.; Dusch, N.; Eggeling, L.; Eikmanns, B.J.; Gaigalat, L. *J Biotechnol*. **2003**, *104*, 5-25.
12. Hermann, T.; Pfefferle, W.; Baumann, C.; Busker, E.; Schaffer, E.; Bott, M.; Sahm, H.; Dusch, N.; Kalinowski, J.; Pühler, A. *Electrophoresis*. **2001**, *22*, 1712-1723.
13. Bendt, A.K.; Burkovski, A.; Schaffer, S.; Bott, M.; Farwick, M.; Hermann, T. *Proteomics*. **2003**, *3*, 1637-1646.
14. Hecker, M. *Adv Biochem Eng Biotechnol*. **2003**, *83*, 57-92.
15. Mader, U.; Hennig, S.; Hecker, M.; Homuth, G. *J Bacteriol*. **2004**, *186*, 2240-2252.
16. Lee, P.S.; Lee, K.H. *Biotechnol Bioeng*. **2003**, *84*, 801-814.
17. Martinez-Antonio, A.; Salgado, H.; Gama-Castro, S.; Gutierrez-Rios, R.M.; Jimenez-Jacinto, V.; Collado-Vides, J. *Biotechnol Bioeng*. **2003**, *84*, 743-749.
18. Tummala, S.B.; Junne, S.G.; Paredes, C.J.; Papoutsakis, E.T. *Biotechnol Bioeng*. **2003**, *84*, 842-854.
19. Lee, P.S.; Shaw, L.B.; Choe, L.H.; Mehra, A.; Hatzimanikatis, V.; Lee, K.H. *Biotechnol Bioeng*. **2003**, *84*, 834-841.
20. Al-Fageeh, M.B.; Marchant, R.J.; Carden, M.J.; Smales ,C.M. *Biotechnol Bioeng*. **2005**, *93*, 829-835.
21. Li, Y.; Powell, S.; Brunette, E.; Lebkowski, J.; Mandalam, R. *Biotechnol Bioeng*. **2005**, *91*, 688-698.

22. Wang, W.; Sun, J.; Hartlep, M.; Deckwer, W.D.; Zeng, A.P. *Biotechnol Bioeng*. **2003**, *83*, 525-536.
23. Piñeiro, C.; Barros-Velázquez, J.; Velázquez, J.; Figueras, A.; Gallardo, J.M. *J Proteome Res*. **2003**, *2,* 127-135.
24. Kern, A.; Tilley, E.; Hunter, I.S.; Legisa, M.; Glieder A. *J Biotechnol*. **2007**, *129*, 6- 29.
25. Wasinger, V.C.; Cordwell, S.J.; Cerpa-Poljak, A.; Yan, J.X.; Gooley, A.A.; Wilkins, M.R.; Duncan, M.W., Harris, R.; Williams, K.L.; Humphery-Smith, I. *Electrophoresis*. **1995**, *7*, 1090-1094.
26. Dutt, M.J.; Lee, K.H. *Curr Opin Biotechnol*. **2000**, *11*, 176-179.
27. Burbaum, J.; Tobal, G.M. *Curr Opin Chem Biol*. **2002**, *6*, 427-433.
28. Butcher, E.C. *Nat Biotechnol*. **2004**, *22*, 1253-1259.
29. Kramer, R.; Cohen, D. *Nat Rev Drug Discov*. **2004**, *3*, 965-972.
30. Onyango, P. *Curr Cancer Drug Targets*. **2004**, *4*, 111-124.
31. Mirza, S.P.; Olivier, M. *Physiol Genomics*. **2008**, *33*, 3-11.
32. Zimmermann, J.D.; Brown, L.R. *Mass Spectrom Rev*. **2001**, *20*, 1-57.
33. Washburn, M.P.; Wolters, D.; Yates, J.R. *Nat Biotechnol*. **2001**, *19*, 242-247.
34. Corthals, G.L.; Wasinger, V.C.; Hochstrasser, D.F.; Sanchez, J.C. *Electrophoresis*. **2000**, *21* (6), 1104-1115.
35. Mayya, V.; Han, D.K. *Expert Review of Proteomics*. **2009**, *6*, 605-618.
36. Cooks, R.G. *Encl Appl Phys*. **1997**, *19*, 289-297.
37. Karas, M.; Hillenkamp, F.; *Anal Chem*. **1988**, *60*, 2299-2301.
38. Whitehouse, C.M.; Dreyer, R.N.; Yamashita, M.; Fenn, J.B. *Anal Chem*. **1985**, *57*, 675-679.
39. Wilm, M.; Mann, M. *Anal Chem*. **1996**, *68*, 1-8.
40. Cornish, T.J.; Cotter, R.J. *Rapid Commun Mass Spectrom*. **1993**, *7*(11), 1037-1040.
41. Dawson, P.H. *Elsevier Scientific Publishing Company*, **1976**.
42. Payne, A.H.; Glish, G.L. *Methods Enzymol*. **2005**, *402*, 109-148.
43. Schwartz, J.C.; Zhou, X.; Bier, M.E. US Patent 5572022, **1996**.
44. Paul, W. *Rev Mod Phys*. **1990**, *62*, 531.
45. Hager, J.W. *Anal Bioanal Chem*. **2004**, *378*, 845-850.
46. Schwartz, J.C.; Senko, M.W.; Syka, J.E.P. *J Am Soc Mass Spectrom*. **2002**, *13*, 659-669.
47. Marshall, A.G.; Hendrickson, C.L.; Jackson, G.S. *Mass Spectrom Rev*. **1998**, *17* (1), 1-35.
48. Domon, B.; Aebersold, R. *Science*. **2006**, *312*, 212-217.
49. McLuckey, S.A.; Goeringer, D.E.; Glish, G.L. *Anal Chem*. **1992**, *64* (13), 1455-1460.
50. Levsen, K. *Verlag Chemie: Weinheim*, **1978**, *138*, 92.
51. Sleno, L.; Volmer, D.A. *J Mass Spectrom*. **2004**, *39* (10), 1091-1112.
52. Paizs, B.; Suhai, S. *Mass Spectrom Rev*. **2005**, *24* (4), 508-548.

53. Hunt, D.F.; Yates, J.R.; Shabanonowitz, J.; Winston, S.; Hauer, C.R. *Proc Natl Acad Sci USA*. **1986**, *83*, 6233-6237.
54. Biemann, K. *Biomed Environ Mass Spectrom*. **1988**, *16*, 99-111.
55. Papayannopoulos, I.A. *Mass Spectrom Rev*. **1995**, *14*; 49-73.
56. Roepstorff, P.; Fohlmann, J. *Biomed Mass Spectrom*. **1984**, *11*, 601.
57. Burlet, O.; Yang, C.Y.; Kaskell, S.J. *J Am Soc Mass Spectrom*. **1992**, *3*, 337-344.
58. Hopfgartner, G.; Varesio, E.; Tschäppät, V.; Grivet, C.; Bourgogne, E.; Leuthold, L.A. *J Mass Spectrom*. **2004**, *39* (8), 845-855.
59. Chernushevich, I.V.; Loboda, A.V.; Thomson, B.A. *J Mass Spectrom*. **2001**, *36*, 849-865.
60. Johnson, J.V.; Jost, R.A.; Kelley, P.A.; Bradford, P.E. *Anal Chem*. **1990**, *62*, 2162-2172.
61. Sabatine, M.S.; Liu, E.; Morrow, D.A.; Heller, E.; McCarroll, R.; Wiegand, R.; Berriz, G.F.; Roth, F.P.; Gerszten, R.E. *Circulation*. **2005**, *112* (25), 3868-3675.
62. Niggeweg, R.; Köcher, T.; Gentzel, M.; Buscaino, A.; Taipale, M.; Akhtar, A.; Wilm, M. *Proteomics*. **2006**, *6* (1), 41-53.
63. Annan, R.S.; Huddleston, M.J.; Verma, R.; Deshaies, R.J.; Carr, S.A. *Anal Chem*. **2001**, *73* (3), 393-404.
64. Douglas, D.J.; Frank, A.J.; Mao D. *Mass Spec Rev*. **2005**, *24*, 1-29.
65. Londrya, F.A.; Hager J.W. *J Am Soc Mass Spectrom*. **2003**, *14* (10), 1130-1147.
66. Hager, J.W. *Rapid Commun Mass Spectrom*. **2002**, *16*, 512-526.
67. Hager, J.W. *Mass spectrometry in proteomics supplement*. **2004**, *3*, 2.
68. Morris, H.R.; Paxton, T.; Dell, A.; Langhorne, J.; Berg, M.; Bordoli, R.S.; Hoyes, J.; Bateman, R.H. *Rapid Commun Mass Spectrom*. **1996**, *10*, 889-896.
69. Shevchenko, A.A.; Chernushevich, I.V.; Ens, W.; Standing, K.G.; Thomson, B.; Wilm, M.; Mann, M. *Rapid Commun Mass Spectrom*. **1997**, *11*, 1015-1024.
70. Douglas, D.J.; French, J.B. *J Am Soc Mass Spectrom*. **1992**, *3*, 398.
71. Krutchinsky, A.N.; Chernushevich, I.V.; Spicer, V.; Ens, W.; Standing, K.G. *J Am Soc Mass Spectrom*. **1998**, *9*, 569.
72. Dodonov, A.F.; Chernushevich, I.V.; Laiko, V.V. *ACS Symp Ser*. **1994**, *549*, 108.
73. Coles, J.; Guilhaus, M. *J Am Soc Mass Spectrom*. **1994**, *5*, 772.
74. Anderson, N.G.; Matheson, A.; Anderson, N.L. *Proteomics*. **2001**, *1*, 3-12.
75. Blackstock, W.P.; Weir, M.P. *Trends Biotechnol*. **1999**, *17*, 121-127.
76. Garfin, D.E. *Trends in Analytical Chemistry*. **2003**, *22*, 263-272.
77. O'Farell, P.H. *J Bio Chem*. **1975**, *CCL*, 4007-4021.
78. Mann, M.; Hendrickson, R.C.; Pandey, A. *Annu Rev Biochem*. **2001**, *70*, 437-473.
79. Rabilloud, T. *Proteomics*. **2002**, *2* (1), 3-10.

80. Unlu, M.; Morgan, M.E.; Minden, J.S. *Electrophoresis*. **1997**, *18* (11), 2071-2077.
81. Gauss, C.; Kalkum, M.; Lowe, M.; Lehrach, H.; Klose, J. *Electrophoresis*. **1999**, *20* (3), 575-600.
82. Berggren, K.N.; Chernokalskaya, E.; Lopez, M.F.; Beechem, J.M.; Patton, W.F. *Proteomics*. **2001**, *1* (1), 54-65.
83. Weckwerth, W.; Willmitzer, L.; Fiehn, O. *Rapid Commun Mass Spectrom*. **2000**, *14*, 1677-1681.
84. Wilm, M.; Mann, M. Anal Chem. **1996**, *68*, 1-8.
85. Wolters, D.A.; Washburn, M.P.; Yates, J.R. *Anal Chem*. **2001**, *73*, 5683-5690.
86. Leitner, A.; Lindner, W. *J Chromatogr B Analyt Technol Biomed Life Sci*. **2004**, *813* (1-2), 1-26.
87. Gygi, S.P.; Rist, B.; Gerber, S.A.; Turecek, F.; Gelb, M.H.; Aebersold, R. *Nat Biotechnol*. **1999**, *17*, 994-999.
88. Fenyo, D.; Qin, J.; Chait, B.T. *Electrophoresis*. **1998**, *19*, 998-1005.
89. Amoresano, A.; Monti, G.; Cirulli, C.; Marino, G. *Rapid Commun Mass Spectrom*. **2006**, *20* (9), 1400-1404.
90. Amoresano, A.; Chiappetta, G.; Pucci, P.; D'Ischia, M.; Marino, G. *Anal Chem*. **2007**, *79* (5), 2109-2117.
91. Palmese, A.; De Rosa, C.; Marino, G.; Amoresano, A. *Rapid Commun Mass Spectrom*. **2011**, *25* (1), 223-231.
92. Ong, S.E.; Foster, L.J.; Mann, M. *Methods* **2003**, *29*, 124-130.
93. Ong, S.E.; Mann, M. *Nat Chem Biol*. **2005**, *1*, 252-262.
94. Steen, H.; Pandey, A. *Trends Biotechnol*. **2002**, *20*, 361-364.
95. Ross, P.L.; Huang, Y.N.; Marchese, J.N.; Williamson, B. *Mol Cell Proteomics*. **2004**, *3*, 1154-1169.
96. Dayon, L.; Hainard, A.; Licker, V.; Turck, N.; Kuhn, K.; Hochstrasser, D.F.; Burkhard, P.R.; Sanchez, J.C. *Anal Chem*. **2008**, *80*, 2921-2931.
97. Bezstarosti, K.; Ghamari, A.; Grosveld, F.G.; Demmers, J.A. *J Proteome Res*. **2010**, *9* (9), 4464-4475.
98. Bonzon-Kulichenko, E.; Pérez-Hernández, D.; Núñez, E.; Martínez-Acedo, P.; Burgos, J.S.; Vázquez, J. *Mol Cell Proteomics*. **2011**, *10* (1), M110.
99. Ong, S.E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D.B. *Mol Cell Proteomics*. **2002**, *1*, 376-386.
100. Tao, W.A., Aebersold, R. *Curr Opin Biotechnol*. **2003**, *14* (1), 110-118.
101. Guevel, L.; Lavoie, J.R.; Perez-Iratxeta, C., Rouger, K.; Dubreil, L.; Feron, M., Talon, S.; Brand, M.; Megeney, L.A. *J Proteome Res*. **2011**, *10* (5), 2465-2478.
102. Krishna, R.G.; Wold F. *In: Creighton T.E., editor. Protein structure: A practical approach. Oxford: Oxford University Press*. **1997**, 91-116.
103. Oda, Y.; Nagasu, T.; Chait, T. *Nat Biotechnol*. **2001**, *19*, 379-382.
104. Nikov, G.; Bhat, V.; Wishnok, J.S.; Tannenbaum, S.R. *Anal Biochem*. **2003**, *320*, 214-222.

105. Zhou, H.; Watts, J.D.; Aebersold, R. *Nat Biotechnol*. **2001**, *19*, 375-378.
106. Lee, S.Y.; Lee, D.Y.; Kim, T.Y. *Trends Biotechnol*. **2005**, *23* (7), 349-358
107. Han, M.J.; Lee, S.Y.; Koh, S.T.; Noh, S.G.; Han, W.H. *J Biotechnol*. **2010**, *145* (4), 341-349.
108. Herrero, E.; Ros, J.; Bellí, G.; Cabiscol, E. *Biochim Biophys Acta*. **2008**, *1780* (11), 1217-1235.
109. Szopinska, A.; Morsomme, P. *OMICS*. **2010**, *14* (6), 639-649.
110. Gessler, N.N.; Aver'yanov, A.A.; Belozerskaya, T.A. *Biochemistry (Mosc)*. **2007**, *72* (10), 1091-1109.
111. Chen, X.; Zhang, W.; Zhang, B.; Zhou, J.; Wang, Y., Yang, Q.; Ke, Y.; He, H. *Proteome Sci*. **2011**, *9*, 37.
112. Bienvenut, W.V.; Espagne, C.; Martinez, A.; Majeran, W.; Valot, B.; Zivy, M.; Vallon, O.; Adam, Z.; Meinnel, T.; Giglione, C. *Proteomics*. **2011**, *11* (9), 1734-1750.
113. Zhan, X.; Desiderio, D.M. *Biochem Biophys Res Commun*. **2004**, *325* (4), 1180-1186.
114. Johnson, H.; Eyers, C.E. *Methods Mol Biol*. **2010**, *658*, 93-108.
115. Manning, G.; Plowman, G.D.; Hunter, T.; Sudarsanam, S. *Trends Biochem Sci*. **2002**, *27*, 514-520.
116. Moorhead, G.B.; De Wever, V.; Templeton, G.; Kerk, D. *Biochem J*. **2009**; *417*, 401-409.
117. Cohen, P. *Trends Biochem Sci*. **2000**, *25*, 596-601.
118. Sefton, B.M.; Shenolikar, S. *Curr Protoc Mol Biol*. **2001**, *Chapter 18* (Unit 18), 11, 3-4.
119. Hemmings, Jr.H.C. *Humana Press*. **1996**, 121-218.
120. Baik, J.Y.; Joo, E.J.; Kim, Y.H.; Lee, G.M. *J Biotechnol*. **2008**, *133*, 461-468.
121. Masaki, S.; Yamada, T.; Hirasawa, T.; Todaka, D.; Kanekatsu, M. *Biotechnol Lett*. **2008**, *30* (5), 955-960.
122. Clarke, S.J.; Khaliulin, I.; Das, M.; Parker, J.E. *Circ Res*. **2008**, *102*, 1082-1090.
123. Chitteti, B.R.; Peng, Z. *Proteomics*. **2007**, *7*, 1473-1500.
124. Kaufmann, H.; Bailey, J.E.; Fussenegger, M. *Proteomics*. **2001**, *1*, 194-199.
125. Eymann, C.; Becher, D.; Bernhardt, J.; Gronau, K. *Proteomics*. **2007**, *7*, 3509-3526.
126. Steen, H.; Kuster, B.; Fernandez, M.; Pandey, A.; Mann, M. *J Biol Chem*. **2002**, *277*, 1031-1039.
127. Graham, M.E.; Anggono, V.; Bache, N.; Larsen, M.R. *J Biol Chem*. **2007**, *282*, 14695-14707.
128. Andersson, L.; Porath, J. *Anal Biochem*. **1986**, *154*, 250-254.
129. Lee, J.; Xu, Y.D.; Chen, Y.; Sprung, R.; Kim, S.C.; Xie, S.H.; Zhao, Y.M. *Mol Cell Proteomics*. **2007**, *6*, 669-676.
130. Ndassa, Y.M.; Orsi, C.; Marto, J.A.; Chen, S.; Ross, M.M. *J Proteome Res*. **2006**, *5*, 2789-2799.

131. Posewitz, M.C.; Tempst, P. *Anal Chem*. **1999**, *71*, 2883-2892.
132. Ficarro, S.B.; McCleland, M.L.; Stukenberg, P.T.; Burke, D.J.; Ross, M.M.; Shabanowitz, J.; Hunt, D.F.; White, F.M. *Nat Biotechnol*. **2002**, *20*, 301-305.
133. Sugiyama, N.; Masuda, T.; Shinoda, K.; Nakamura, A.; Tomita, M.; Ishihama, Y. *Mol Cell Proteomics*. **2007**, *6*, 1103–1109.
134. Jensen, S.S.; Larsen, M.R. *Rapid Commun Mass Spec*. **2007**, *21*, 3635-3645.
135. Wang, J.L.; Zhang, Y.J.; Jiang, H.; Cai, Y.; Qian, X.H. *Proteomics*. **2006**, *6*, 404-411.
136. Pinkse, M.W.; Mohammed, S.; Gouw, L.W.; van Breukelen, B.; Vos, H.R.; Heck, A.J. *J Proteome Res*. **2008**, *7*, 687-697.
137. Cantin, G.T.; Shock, T.R.; Park, S.K.; Madhani, H.D.; Yates, J.R. *Anal Chem*. **2007**, *79*, 4666-4673.
138. Lim, K.B.; Kassel, D.B. *Anal Biochem*. **2006**, *354*, 213-219.
139. Chen, M.; Su, X.; Yang, J.; Jenkins, C.M.; Cedars, A.M.; Gross, R.W. *Anal Chem*. **2010**, *82* (1), 163-171.
140. Zhou, W.; Merrick, B.A.; Khaledi, M.G.; Tomer, K.B. *J Am Soc Mass Spectrom*. **2000**, *11* (4), 273-282.
141. Li, W.; Boykins, R.A.; Backlund, P.S.; Wang, G.; Chen, H.C. *Anal Chem*. **2002**, *74* (22), 5701-5710.
142. Adamczyk, M.; Gebler, J.C.; Wu, J. *Rapid Commun Mass Spectrom*. **2001**, *15* (16), 1481-1488.
143. Poulter, L.; Ang, S.G.; Gibson, B.W.; Williams, D.H.; Holmes, C.F.; Caudwell, F.B.; Pitcher, J.; Cohen, P. *Eur J Biochem*. **1988**, *175* (3), 497-510.
144. Ohguro, H.; Palczewski, K.; Ericsson, L.H.; Walsh, K.A.; Johnson, R.S. *Biochemistry*. **1993**, *32* (21), 5718-5724.
145. Huddleston, M.J.; Annan, R.S.; Bean, M.F.; Carr, S.A. *J Am Soc Mass Spectrom* **1993**, *4*, 710-717.
146. Hunter, A.P.; Games, D.E. *Rapid Commun Mass Spectrom*. **1994**, *8*, 559-570.
147. Schlosser, A.; Pipkorn, R.; Bossemeyer, D.; Lehmann, W.D. *Anal Chem*. **2001**, *73*, 170-176.
148. Galley, H.F. *Br J Anaesth*. **2011**, *107* (1), 57-64.
149. Dalle Donne, I.; Scaloni, A.; Giustarini, D.; Cavarra, E.; Tell, G.; Lungarella, G.; Colombo, R.; Rossi, R.; Milzani, A. *Mass Spectrom Rev*. **2005**, *24* (1), 55-99.
150. Murphy, M.P. *Biochem J*. **2009**, *417*, 1-13.
151. Turrens, J.F. *J Physiol*. **2003**, *552*, 335-344.
152. Lee, M.Y.; Griendling, K.K.; *Antioxid. Redox Signal*. **2008**, *10* (6), 1045-1059.
153. Calabrese, V.; Guagliano, E.; Sapienza, M.; Mancuso, C.; Butterfield, D.A.; Stella, A.M. *Ital J Biochem*. **2006**, *55* (3-4), 263-282.
154. Fialkow, L.; Wang, Y.; Downey, G.P. *Free Radic Biol Med*. **2007**, *42* (2), 153-164.
155. Sies, H. *Exp Physiol*. **1997**, *82* (2), 291-295.

156. McMahon, M.; Grossman, J.; FitzGerald, J.; Dahlin-Lee, E.; Wallace, D.J.; Thong, B.Y.; Badsha, H.; Kalunian, K.; Charles, C.; Navab, M.; Fogelman, A.M.; Hahn, B.H. *Arthritis Rheum*. **2006**, *54* (8), 2541-2549.
157. Emerit, J.; Edeas, M.; Bricaire F. *Biomed Pharmacother*. **2004**, *58* (1), 39-46.
158. Matés, J.M.; Segura, J.A.; Alonso, F.J.; Márquez, J.; *Arch Toxicol*. **2008**, *82* (5), 273-299.
159. Bannister, J.V.; Bannister, W.H.; Rotilio, G.; *CRC Crit Rev Biochem*. **1987**, *22* (2), 111-180.
160. Afanas'ev, I.B. *Mol Biotechnol*. **2007**, *37* (1), 227-478.
161. Lundberg, J.O.; Weitzberg, E.; Gladwin, M.T. *Nat Rev Drug Discov*. **2008**, *7* (2), 156-167.
162. Murad, F. *Biosci Rep*. **2004**, *24* (4-5), 452-474.
163. Hughes, M.N. *Methods Enzymol*. **2008**, *436*, 3-19.
164. Beckman, J.S.; Koppenol, W.H. *Am J Physiol*. **1996**, *271* (5), C1424-C1437.
165. Landolfo, S.; Politi, H.; Angelozzi, D.; Mannazzu, I. *Biochim Biophys Acta*. **2008**, *1780* (6), 892-898.
166. Ischiropoulos, H. *Arch Biochem Biophys*. **1998**, *356* (1), 1-11.
167. Schopfer, F.J.; Baker, P.R.; Freeman, B.A. *Trends Biochem Sci*. **2003**, *28* (12), 646-654.
168. Radi, R. *Proc Natl Acad Sci USA*. **2004**, *101* (12), 4003-4008.
169. Pfeiffer, S.; Schmidt, K.; Mayer, B. *J Biol Chem*. **2000**, *275* (9), 6346-6352.
170. De Caro, J.D.; Behnke, W.D.; Bonicel, J.J.; Desnuelle, P.A.; Rovery, M. *Biochim Biophys Acta*. **1983**, *3*, 253-262.
171. Ischiropoulos, H.; Zhu, L.; Chen, J.; Tsai, M.; Martin, J.C.; Smith, C.D.; Beckman, J.S. *Arch Biochem Biophys*. **1992**, *298*, 431-437.
172. Souza, J.M.; Daikhin, E.; Yudkoff, M.; Raman, C.S.; Ischiropoulos, H. *Arch Biochem Biophys*. **1999**, *371*, 169-178.
173. Greenacre, S.A.; Ischiropoulos, H. *Free Radic Res*. **2001**, *34*, 541-581.
174. Ohshima, H.; Friesen, M.; Brouet, I.; Bartsch, H. *Food Chem Toxicol*. **1990**, *28* (9), 647-652.
175. Aulak, K.S.; Koeck, T.; Crabb, J.W.; Stuehr, D.J. *Methods Mol Biol*. **2004**, *279*, 151-165.
176. Butt, Y.K.; Lo, S.C. *Methods Enzymol*. **2008**, *440*, 17-31.
177. Cueni, L.; Riordan, J.F. *Biochemistry*. **1978**, *17* (10), 1834-1842.
178. Muszynska, G.; Riordan, J.F. *Biochemistry*. **1976**, *15* (1), 46-51.
179. Kanski, J.; Schoneich, C. *Methods Enzymol*. **2005**, *396*, 160-171.
180. Bigelow, D.J.; Qian, W.J. *Methods Enzymol*. **2008**, *440*, 191-205.
181. Sacksteder, C.A.; Qian, W.J.; Knyushko, T.V.; Wang, H.; Chin, M.H.; Lacan, G.; Melega, W.P.; Camp, D.G.; Smith, R.D.; Smith, D.J.; Squier, T.C.; Bigelow, D.J. *Biochemistry*. **2006**, *45* (26), 8009-8022.
182. Sarver, A.; Scheffler, N.K.; Shetlar, M.D.; Gibson, B.W. *J Am Soc Mass Spectrom*. **2001**, *12* (4), 439-448.

183. Petersson, A.S.; Steen, H.; Kalume, D.E.; Caidahl, K.; Roepstorff, P. *J Mass Spectrom*. **2001**, *36* (6), 616-625.
184. Petre, B.A.; Youhnovski, N.; Lukkari, J.; Weber, R.; Przybylski, M. *Eur J Mass Spectrom*. **2005**, *11* (5), 513-518.
185. Turko, I.V.; Murad, F. *Methods Enzymol*. **2005**, *396*, 266-275.
186. Zhan, X.; Desiderio, D.M. *Int J Mass Spectrom*. **2007**, *259* (1-3), 96-104.
187. Davies, K.J.; Delsignore, M.E. *J Biol Chem*. **1987**, *262* (20), 9895-9901.
188. Davies, K.J.; Delsignore, M.E. *J Biol Chem*. **1987**, *262* (20), 9914-9920.
189. Davies, K.J.; Delsignore, M.E.; Lin, S.W. *J Biol Chem*. **1987**, *262* (20), 9908-9913.
190. Davies, K.J. *J Biol Chem*. **1987**, *262* (20), 9902-9907.
191. Stadtman, E.R.; Berlett, B.S. *Drug Metab Rev*. **1998**, *30*, 225–243.
192. Berlett, B.S.; Stadtman, E.R.; *J Biol Chem*. **1997**, *272*, 20313–20316.
193. Amici, A.; Levine, R.L.; Tsai, L.; Stadtman, E.R.; *J Biol Chem*. **1989**, *264*, 3341-3346.
194. Requena, J.R.; Stadtman, E.R. *Biochem Biophys Res Commun*. **1999**, *264*, 207-211.
195. Miyata, T.; Inagi, R.; Asahi, K.; Yamada, Y.; Horie, K.; Sakai, H.; Uchida, K.; Kurokawa, K.; *FEBS Lett*. **1998**, *437*, 24-28.
196. Morgan, P.E.; Dean, R.T.; Davies, M.J. *Arch Biochem Biophys*. **2002**, *403*, 259-269.
197. Levine, R.L. *J Biol Chem*. **1983**, *258*, 11828-11833.
198. Stadtman, E.R. *Free Rad Biol Med*. **1990**, *9*, 315-325.
199. Uchida, K.; Kato, Y.; Kawakishi, S. *Biochem Biophys Res Commun*. **1990**, *169* (1), 265-271.
200. Garrison, W.M. *Chem Rev*. **1987**, *87*, 381-398.
201. Lee, A.T.; Cerami, A. *In: Schneider EL, Rowe JW (eds) Handbook of the biology of aging, 3rd edn. Academic Press, New York*. **1990**, 116-130.
202. Verzijl, N.; De Groot, J.; Oldehinkel, E.; Bank, R.A.; Thorpe, S.R.; Baynes, J.W.; Bayliss, M.T.; Bijlsma, W.J.; Lafeber, F.P.; Tekoppele, J.M. *Biochem J*. **2000**, *350*, 381-387.
203. Wolf, S.P.; Garner, A.; Dean, R.T. *Trends Biochem Sci*. **1986**, *11*; 1-5.
204. Uchida, K.; Stadtman, E.R.; *J Biol Chem*. **1993**, *268*, 6388-6393.
205. Levine, R.L. *Free Radic Biol Med*. **2002**, *32*, 790-796.
206. Dukan, S.; Nystrom, T. *Genes Dev*. **1998**, *12*, 3431-3441.
207. Cabiscol, E.; Piulats, E.; Echave, P.; Herrero, E.; Ros, J. *J Biol Chem*. **2000**, *275*, 27393–27398.
208. Sohal, R.S. *Free Radic Biol Med*. **2002**, *33*, 37–44.
209. Nystrom, T. *Embo J*. **2005**, *24*, 1311–1317.
210. Lee, S.; Young, N.L.; Whetstone, P.A.; Cheal, S.M.; Benner, W.H. *J Proteome Res*. **2006**, *5*, 539–547.
211. Mirzaei, H.; Regnier, F. *Anal Chem*. **2005**, *77*, 2386–2392.
212. Mirzaei, H.; Regnier, F. *Anal Chem*. **2006**, *78*, 770–778.

213. Mirzaei, H.; Regnier, F. *J Proteome Res*. **2006**, *5*, 2159–2168.
214. Temple, A.; Yen, T.Y.; Gronert, S. *J Am Soc Mass Spectrom*. **2006**, *17*, 1172–1180.
215. Dalle Donne, I.; Giustarini, D.; Colombo, R.; Rossi, R.; Milzani, A. *Trends Mol Med*. **2003**, *9* (4), 169-176.
216. Choi, J.; Malakowsky, C.A.; Talent, J.M.; Conrad, C.C.; Gracy, R.W. *Biochem Biophys Res Commun*. **2002**, *293*, 1566-1570.
217. Lenz, A.G.; Costabel, U.; Shaltiel, S.; Levine, R.L. *Anal Biochem*. **1989**, *177*, 419-425.
218. Levine, R.L.; Wehr, N.; Williams, J.A.; Stadtman, E.R.; Shacter, E. *Methods Mol Biol*. **2000**, *99*, 15-24.
219. Climent, I.; Tsai, L.; Levine, R.L. *Anal Biochem*. **1989**, *182*, 226-232.
220. Robinson, C.E.; Keshavarzian, A.; Pasco, D.S.; Frommel, T.O.; Winship, D.H.; Holmes, E.W. *Anal Biochem*. **1999**, *266*, 48-57.

II. QUANTITATIVE IDENTIFICATION OF PROTEIN NITRATION SITES

II.1 Introduction

In this chapter I report a novel proteomics strategy, based on the use of iTRAQ reagents coupled to mass spectrometry analysis, for the selective analysis of *o*-nitrotyrosine residues. This method was proved to lead the simultaneous localisation and quantification of nitration sites both in model proteins and in biological systems. The strategy proposed may be classified as belonging to a class of general derivatization methods defined with the acronym RIGhT (Reporter Ion Generating Tag) (see chapter I.3.1). This approach is based on the selective labelling of the target residues with reagents capable to generate reporter ions in MS^2 experiments. The detection of the reporter is a very stringent criterion that is used to reduce the analytical complexity to a few numbers of species, resulting in an improved duty cycle and enhanced dynamic range achievable by the MS analysis. In operative terms, the use of a high selective MS replaces the off-line chromatographic steps (affinity and/or bidimensional chromatography) aimed to enrich the mixture in labelled peptides. So the strategy proposed in this chapter can be considered as a *third generation* proteomics methodology.

II.1.1 Proteomics analysis of nitrated proteins

A number of so-called tagging (or labelling) strategies have been developed to target specific amino acid residues and/or post-translational modifications (PTMs). These methods enable the enrichment of subfractions via affinity clean-up. A practical approach consists in targeting the specific structural traits followed by an affinity chromatography step [1-3]. This approach simplifies the proteome while preserving most of the vital information necessary for the analysis. With a similar strategy, it is also possible to isolate post-translationally modified peptides [4-6]. The presence of phosphorylation, glycosylation, glycation, nitration, and specific types of oxidation are examples of PTMs that can be targeted. Analysis of PTMs provides one of the strongest cases for targeted selection. Although many proteins are post-translationally modified during their biological lifetime, perhaps only one out of 20-50 of their tryptic peptides will be modified, depending on the PTM. Thus after a trypsin digestion only one or few peptides carrying the modifying group will be generated. As a consequence, PTM specific selection will lead to a number of peptides to be examined reduced at least of an order of magnitude [7].
Several modifications have been described and they differ substantially in the extent of the modification of proteins. For this reason, some PTMs can be targeted more specifically than others, and the tagging concepts currently described in the literature are most often directed towards phosphorylated proteins, although recently methods for the study of glycosylation and tyrosine nitration have been proposed [8-9].
Among PTMs, protein nitration can occur in cells during oxidative stress and over-production of nitric oxide (section I.3.2.1) [10-11]. Nitration sites in proteins are phenylalanine, tryptophan and, in particular, tyrosine residues. Tyrosine nitration is becoming increasingly recognized as a prevalent, functionally significant protein post-translational modification. Nitration of proteins modulates catalytic activity, cell signalling and cytoskeletal organisation [12]. Addition of NO_2 group to the orto position of tyrosine decreases the pK_a of the tyrosine hydroxyl group of about three units; this

introduces a net negative charge to approximately half of nitrated tyrosine residues at physiological pH. This bulky substituent can induce changes in protein conformation, resulting in the generation of antigen epitopes, altered enzyme catalytic activity, modulation of metabolic pathways and inhibition of tyrosine phosphorylation by protein kinases [13-15]. For these reasons and for its chemical stability NO_2Tyr is considered the most important biomarker for identification and quantitation of cellular processes, associated to RNS over-production, that lead post-translational modification of proteins.

Many research groups have focused their attention towards 3-nitrotyrosine residues, developing various methodologies for revelation and quantification of these residues. Most of the studies published so far focused on the detection and quantification of *o*-nitrotyrosine by the use of 2D-GE and Western blot, followed by the MS analysis of the *in situ* digested proteins, positive to the immuno-assay. For 2D-GE fractionated proteins, the identification based on the fragmentation spectra were presented in only a limited number of works [16-17]. This lacking of data may be caused by the classical limitations of the 2D-GE technique (section 1.2.2). In particular the low quantity of proteins analysable by IEF is very limiting considering the sub-stoichiometric nature of most PTMs. Immune-histochemical or ELISA techniques have been used as alternative to 2D-GE and Western blot but are unable to specifically identify the site of modification. Moreover these procedures generally led to the identification of nitrated proteins but the exact localization of the nitration sites in complex protein mixtures still remain a challenging task. More recently chromatographic methods have been developed using either high performance liquid chromatography (HPLC) separation combined with electrochemical detection or GC-MS techniques [18-20]. In addition at the common 2D-GE limits, also the MS and MS^2 analysis of nitrated peptides has some technical aspects that have to be considered. The development of gel free methods has represented an advance in nitroproteins analysis, allowing to work with a much larger amount of nitrated proteins or peptides. By using multidimensional chromatography-tandem mass spectrometry (SCX-RP-MS^2) approaches, the rat neuroproteome was explored identifying 31 sites of nitration in 29 different proteins [21]. However, although this approach allows to overcome many limits of gel-based methods, the detection of lower abundant nitrated proteins remains problematic. To analyse these PTMs at a proteomics level, a specific enrichment step is highly advised prior to MS analysis. The unique method to directly enrich nitro-proteins uses anti-nitrotyrosine based affinity chromatography. However this strategy is not so spread [22-23]. This is probably due to the low specificity of anti-nitrotyrosine antibodies. There are no effective enrichment methods for nitrated peptides, probably because of the poor chemical reactivity of nitro group. However, different strategies were introduced based on the conversion of NO_2Tyr nitro group in amino group, which is much more reactive. Nikov *et al*. reported a novel method for the identification of peptides containing nitrotyrosine residues in complex mixture [24]. This strategy involved first free thiols blocking by iodoacetamide, then nitrotyrosine reduction to aminotyrosine. Then aminotyrosines residues were labelled with a cleavable biotin tag and the protein sample was digested. The biotinylated peptides were purified on a streptavidin derivatized solid support. Further nitrated peptides were eluted and they were directly analysed by MALDI-TOF mass spectrometry. Zhang *et al.* introduced an improved enrichment strategy [25]. In fact they derivatized reduced aminotyrosines with N-succinimidyl S-acetylthioacetate (SATA). Then the thio-esther group was converted to thiol allowing purification by using thiopropyl sepharose beads. The enriched peptides, analyzed by LC-MS/MS, showed 6-10

times more nitrotyrosine-containing peptides than the global analysis of not enriched samples.
MALDI mass spectrometry has been extensively used for the analysis of nitrotyrosines; although NO_2Tyr residues formation causes an increment of 45 Da of the molecular mass of the modified peptide respect to the unmodified one, nitrotyrosine containing peptides are sensitive to laser light and undergo a decomposition process yielding, besides the expected mass increase of 45 Da, other major peaks corresponding to nitrosotyrosine ((Tyr-NO) [M + H^+ + 29]), aminotyrosine ((Tyr-NH_2) [M + H^+ + 15]), and nitrenetyrosine ((Tyr-N) [M + H^+ +13]) [26-27]. While on the one hand this phenomenon constitutes a unique fingerprint, it also impairs the sensitivity of the analysis, since it spreads the nitrated peptide signal over at least three major species, hence reducing the signal intensity by as much as 60-70%, which may result in the failure to observe nitrated peptides in complex peptide mixtures [28-30]. This is worsened by the fact that tyrosine nitration is a low-abundance PTM with an occurrence of 1 in 10000 tyrosines *in vivo* [31].
It has been showed by Petersson and co-workers [27] that using ESI mass spectrometry, this phenomenon is absent. Moreover it is possible to use ESI-MS/MS in precursor ion scan mode to detect nitrated peptides in complex mixtures using a characteristic fragment ion of nitrated peptides, such as the immonium ion of nitrotyrosine [27].
My tutor's research group has recently suggested a strategy leading to the selective detection of nitrated tyrosine residues by using dansyl chloride labelling coupled to advanced mass spectrometry experiments [32]. The tryptic digest from the entire protein mixture is directly analyzed by MS on a linear ion trap mass spectrometer. Discrimination between nitro- and unmodified peptide is based on two selectivity criteria obtained by combining a precursor ion scan and an MS^3 analysis. The proposed strategy is certainly rapid and easy to use but the major limit resulted in the absence of heavy form of dansyl chloride. Thus quantitative analysis of complex protein mixtures cannot be addressed. Taking advantage of the above mentioned strategy, i used selective labelling of 3-nitrotyrosine residues exploiting the potential of iTRAQ reagents to perform quantitative analysis of this post-translational modification in proteins. The strategy proposed is outlined in the figure II.1.

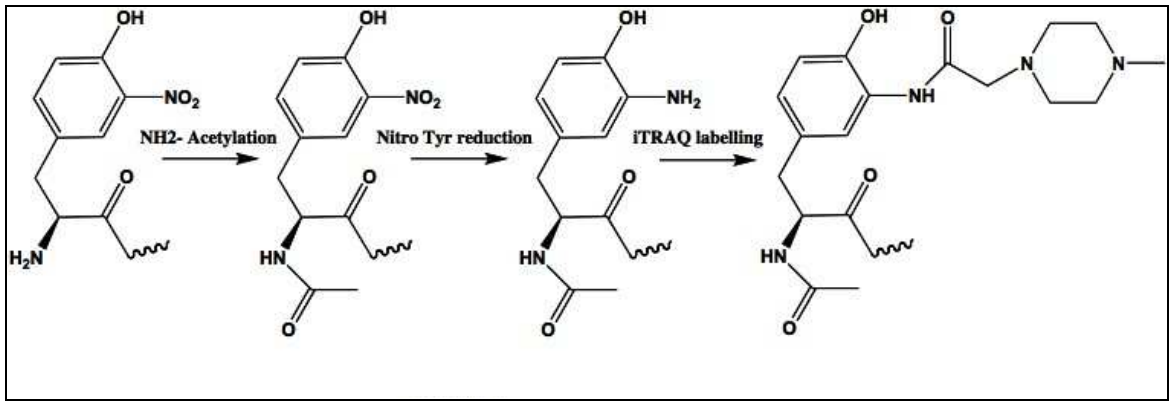


Figure II.1. *Strategy for selective labelling of 3-nitrotyrosine residues*

II.1.2 Quantitative analysis for nitrated proteins

So far the iTRAQ chemistry has been limited to primary amines. With my work I am planning to widen the chemistry of this reagent to address quantification of functions

other than primary amines, by taking advantage of the experience made with dansyl chemistry. In fact, the research group in which I worked for my PhD thesis has already reported a totally new approach involving dansyl chloride labelling of the nitration sites that rely on the enormous potential of MS^n analysis [32].
Stable isotope methods have been introduced into MS-based proteomics to allow to determine relative changes in protein expression [33]. The principle of these methods is the incorporation of a stable isotope derivative in one of the states to be compared. Stable isotope incorporation shifts the mass of the peptides by a predictable amount. The ratio of analyte between the two or more states can then be determined accurately by the measured peak ratio between the heavy and light derivatized samples. One of the first approaches based on the isotope stable affinity tagging and MS is the isotope coded affinity tag strategy (ICAT) [34] specifically addressed towards Cys residues. Other elegant approach uses cell culture enrichment with a stable isotope-labelled amino acid, for *in vivo* incorporation of a mass difference to perform relative quantitation [34].
A novel methodology for quantitative analysis by mass spectrometry makes use of iTRAQ (acronym for Isobaric Tag for Relative and Absolute Quantification) technology, a newly developed method by Applied Biosystems for relative and absolute quantification of proteins [35]. The iTRAQ reagents are specifically reactive towards primary amino groups (namely N-terminal of proteins and peptides and ε-amino group of lysine residues) and marketed in four different forms called iTRAQ 114, 115, 116 and 117, depending of m/z value of reporter group. The reagents are differentially isotopically labelled such that all derivatized peptides are isobaric (isobaric tag with a mass of 145.1 Da). In fact, 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. The iTRAQ reagents allow the simultaneously multiplexed analysis of four or eight samples. Each of these samples is labelled with one of the iTRAQ reagents; all peptides with the same sequence, but carrying different versions of the tag will be chromatographically indistinguishable and identical in mass, therefore, also identical in single MS mode. Upon fragmentation, the reporter group is detached, creating signals in the “quite” low mass range region (114-117 Da) of the MS/MS spectrum. The balance group is lost as a neutral, so that the remaining peptide backbone remains unmodified and can generate backbone fragments that are identical in m/z for all samples, resulting in improved S/N ratios because all differentially coded samples contribute to MS/MS spectrum. Thus low mass reporter ion signals allow quantitation, while peptide fragment ion signals allow protein identification. Very recently, an octu-plex iTRAQ kit has been proposed thus allowing the simultaneous analysis of eight different samples [36].

II.1.2 iTRAQ as a RIGhT reagent

RIGhT approach is based on the selective MS analysis of species of interest by using the molecular structure features of labelled peptides and the MS^2 characteristic fragmentation pathway introduced by the chosen reagent. RIGhT reagents have a peculiar structure consisting in a reactive group toward the interesting residues, a linker group that is loss as neutral moiety in MS^2 experiment and a reporter group that leads to generate stable fragment ions, selectively detectable in MS^2 experiment. For the development of a RIGhT strategy that allows the identification and quantitation of nitrated proteins, iTRAQ was chosen as RIGhT reagent. iTRAQ is a

primary amines reactive molecule, with all the structural features of a RIGhT reagent (figure II.2).

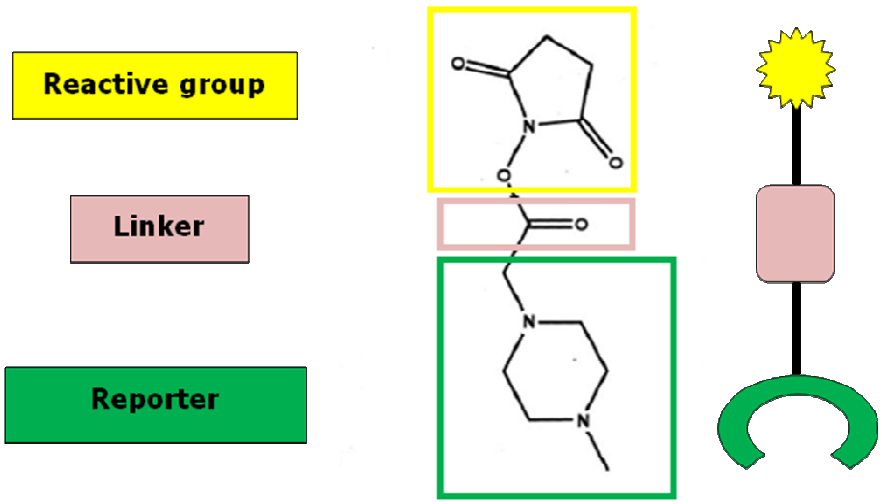


Figure II.2. *iTRAQ as RIGhT reagent.*

After MS^2 fragmentation, iTRAQ peptides give rise to peculiar ions with m/z 114-117 according to the iTRAQ molecule used for labelling. These characteristics fragments were exploited by using an Applied Biosystems 4000QTrap mass spectrometer. By the coupling of this instrumentation and of the iTRAQ derivatives characteristics, it was set up a high throughput LC-MS/MS method based on *Precursor Ion Scan* (PIS) analysis. PIS was chosen as survey scan for its selectivity and its high scan rate. In this scan mode the Q_1 performs a scan across the full mass range; the ions separated are fragmented in the collision cell (q_2); finally the Q_3 is set to transmit only ions with the specific mass of the diagnostic fragment (m/z 114-117). Selected precursor ions are then submitted to a classical MS^2 experiment exploiting the LIT features of the QTrap. This MS based selection of labelled peptides substitutes others chromatographic steps so that the mass spectrometer can be considered as the second dimension of the RP-LC separation.

II.2 Materials and methods

Chemicals

Bovine serum albumin (BSA), tetranitromethane (TNM), ammonium bicarbonate (AMBIC), guanidine, dithiothreitol (DTT), trypsin, iodoacetamide (IAM) and triethyl ammonium carbonate (TEAB) were purchased from Fluka. Tris(hydroxymethyl)aminomethane, sodium dithionite as well as the MALDI matrix α-ciano-4-idrossicinnamic acid were purchased from Sigma. iTRAQ reagents were furnished by Applied Biosystems. Methanol, trifluoroacetic acid (TFA) and acetonitrile (ACN) are HPLC grade type from Carlo Erba, whereas the other solvents are from Baker. Gel filtration columns PD-10 are from Pharmacia, the HPLC ones from Phenomenex, whereas the pre-packed columns Sep-pak C-18 are from Waters.

BSA nitration

An aliquot of BSA was dissolved in Tris 200 mM pH 8.0 buffer and was added an 830 nmol/ml solution of TNM in ACN in molar ratio 1:10 BSA : TNM. Reaction was carried out under agitation for 30 min at room temperature. Product was then purified by size

exclusion chromatography on a SEPHADEX G-25M column; the elution was carried out with Tris 300 mM and the fractions analyzed through UV spectrometry at 220 and 280 nm wavelength. Positive fractions were collected and dried.

Reduction and carboamidomethylation of BSA
Dried nitrated BSA (N-BSA) was dissolved in denaturation buffer (guanidine 6 M, Tris 0.3 M, EDTA 10 mM, pH 8.0). Reduction was carried out by using a 10:1 DTT:cysteine molar ratio. After incubation at 37℃ for 2 h, iodoacetamide was added to perform carboamidomethylation using an excess of alkylating agent of 5:1 respect to the moles of thiolic groups. The mixture was then incubated in the dark at room temperature for 30 min. The alkylation reaction was stopped by addition of formic acid, in order to achieve an acidic pH. The product was purified by size exclusion chromatography. The elution was performed with TEAB 50 mM pH 8.0.

BSA digestion
Digestion of N-BSA was carried out in TEAB 50mM pH 8.0 buffer using trypsin at a 50:1 protein : trypsin mass ratio. The sample was incubated at 37℃ for 16 h. Then sample was dried.

Acetylation of peptide mixture
Peptide mixture was dissolved in 30% TEAB 500mM pH 8.0, 70% ACN. Acetylation was performed by using acetic acid N-hydroxy-succinimide esther in molar ratio 500:1 reagent : peptides. The reaction was carried out for 1 h at room temperature.

Nitro groups reduction
Reduction of nitro groups to amino groups was performed by adding. $Na_2S_2O_4$ at 100:1 $Na_2S_2O_4$: NO_2Tyr molar ratio. Reaction was carried out at room temperature for 10 min under stable stir. Product was purified by SEP-PAK RP-LC with C18 column. Peptides were eluted using 80% ACN, 20% formic acid 0.1%. Fractions were collected and lyophilized.

iTRAQ selective labelling
Sample was dissolved in TEAB 500mM pH 8.0 and then was divided in two shares and each of these was differentially labelled with alternatively iTRAQ reagent 114 and iTRAQ reagent 117, following a modified version of iTRAQ labelling protocol furnished by Applied Biosystems; the reaction was carried out in 60% ACN, 40% TEAB 500mM pH 8.0, for 2 h at 37 ℃.

Labelling of Bovine Milk Protein Extract
A sample of commercially available bovine milk was reacted with a 10 mM solution of TNM for 30 min at room temperature. Modified milk proteins were purified by precipitation with the Amersham Clean Up kit and dissolved in denaturant buffer. Reduction of both the SH and nitro groups of the protein mixture was carried out in “one pot” using 20 mM DTT, 40 mM iodoacetamide, and a 200 mM solution of $Na_2S_2O_4$ in the dark for 30 min. The protein mixture was purified by size exclusion chromatography on a Sephadex G-25 M column equilibrated and eluted with 50 mM TEAB. Protein fractions were concentrated and then digested with trypsin as already described. The resulting peptide mixture was acetylated and reacted with ITRAQ as described above.

nanoLC Mass Spectrometry

Peptide mixture, obtained as previously described, was analysed by LC-MS/MS analysis using a 4000QTrap (Applied Biosystems) coupled to an 1100 nanoHPLC 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). Nanospray source was used at 2.3 kV with liquid coupling, with a declustering potential of 20 eV, using an uncoated silica tip from NewObjectives (O.D. 150 μm, I.D. 20 μm, T.D. 10 μm). Spectra acquisition was based on a survey *precursor ion scanning* for the ion m/z 114 and 117. It was performed over a mass range of m/z 400-1400 with Q_1 set to low resolution and Q_3 set to unit resolution. Precursors were collided in q_2 with a collision energy ramp from 25 to 65 eV across the mass range. Spectra acquisition was based on a survey precursor ion scan. A positive ion *enhanced resolution scan* was performed at 250amu/s to determine the charge state of the ion. *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 and the cycle repeated. The entire cycle duration, including fill times and processing times was less than 4.3 s. Data were acquired using Analyst software (Applied Biosystems) and processed by using in house Mascot software.

II.3 Results and discussions

According to the so-called “gel-free procedures”, the analysis is carried out at level of peptides following tryptic digest of the whole protein mixture. The approach presented in this chapter was first applied to bovine serum albumin (BSA) as model protein to prove its effectiveness. The BSA was nitrated using TNM as described in materials and methods section. After nitration, aliquots of BSA and N-BSA mixtures were reduced, alkylated and hydrolyzed using trypsin. The peptide mixtures were then analysed via MALDI-MS to verify the extent of nitration.

II.3.1 Analysis of 3-nitrotyrosine residues of a model protein

BSA is a protein of about 66 kDa and it has 20 tyrosine residues along its sequence. However nitration was limited to only few tyrosine residues, namely tyrosine residues exposed in the structure of protein and thus more sensible to nitration.

Nitrated peptides were identified by MALDI *mapping* procedure, by comparing the experimental peptides masses with the theoretical values on the basis of amino acid sequence. As already reported a ΔM of +45 was attributed to the modified peptide, due to introduction of a nitro group [37]. The MALDI-MS spectrum showed the occurrence of four signals at m/z 972.5, 1524.8, 1612.9, 1770.8, respectively whose molecular masses were 45 Da higher than the theoretical ones. These signals were attributed to the peptides 143-149, 403-415, 329-341, 451-464 having the Tyr residues modified by a nitro group. As further proof, these signals exhibited the characteristic pattern of nitrated peptides, due to photodecomposition of nitrated tyrosine residues by MALDI source [37] (figure II.3).

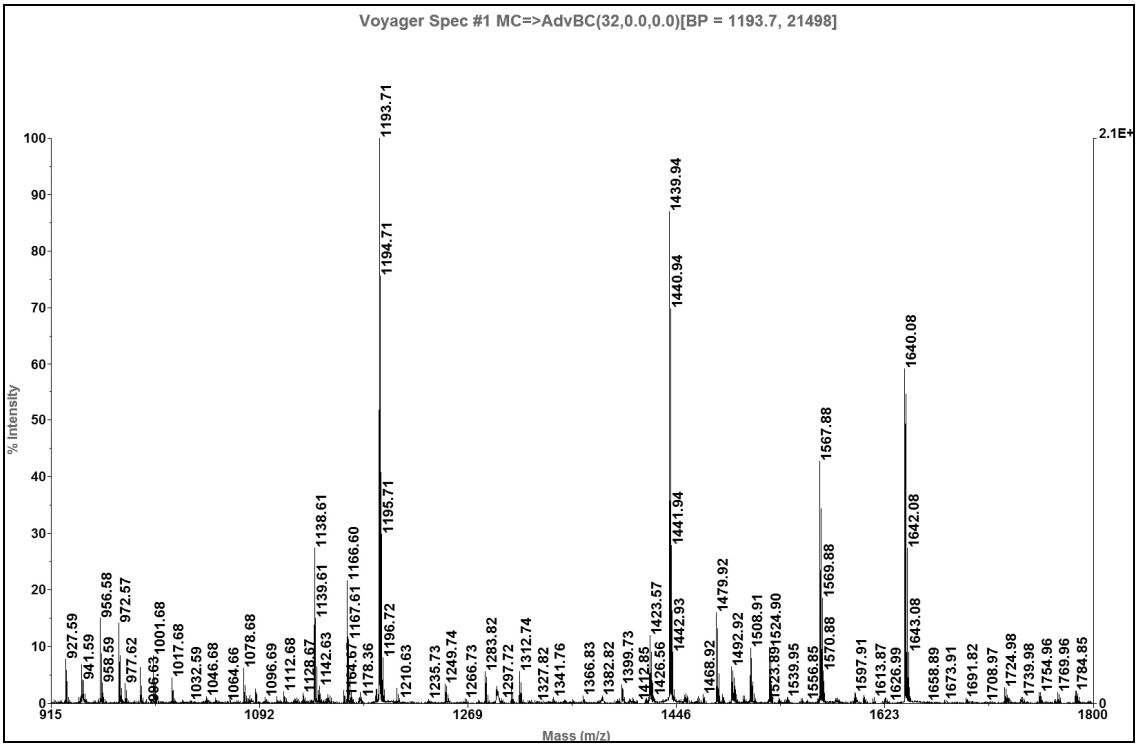


Figure II.3. *Detail of a MALDI spectrum obtained from the analysis of nitrated BSA. The spectrum shows the presence of a signal at m/z 972.57 attributable to the nitrated peptide 143-149 and the peaks at m/z 956.58, 941.59 because of the photodecomposition on nitrated tyrosine residue.*

In our previous work [32], a new methodology for the selective labelling of *o*-nitrotyrosine residues was essentially based on a reduction step and conversion to *o*-aminotyrosine residues followed by specific dansylation at pH 4.7. Similar strategy was attempted here. However, iTRAQ labelling does not permit to exploit the characteristic pKa 4.7 [38] of such aromatic amine to obtain a selective labelling, as proposed for the dansyl chloride labelling. In fact, although in solution at pH 5.0 the free amino groups able to give nucleophilic attack, are not abundant, a small percentage of primary amine modification was observed, as also already demonstrated [39-40]. This behaviour may be due to an "ankymeric effect", namely an increased reactivity of iTRAQ molecule due to both slower rate of hydrolysis of the N-hydroxysuccynimide esther in mild acidic condition and the acid activation of acylic nucleophilic substitution [41]. Thus, a double step strategy was set up. First of all, lysine residues and N-terminus primary amines were protected to facilitate specific labelling of the 3-nitrotyrosine residues. Blocking can be easily performed under mild conditions by a variety of reagents that are commercially available. In fact, several strategies to introduce NH_2- labels have been described [42-43]. Generally, it is not possible to reliably control the selectivity of the reaction, because the ε-amino group of lysine is readily and stably modified by reagents that target the N-terminus [44-45]. Blocking the N-terminus and the lysine side chain can be done either prior to or after enzymatic digestion of the sample. By blocking the lysines after digestion, we retained the ability to use trypsin as enzyme for proteolytic cleavage. Thus, nitrated peptides were subjected to acetylation using acetic acid N-hydroxy-succinimide esther as described in material and methods section. All successive analyses revealed that every peptide detected had the N-terminal amine and eventually ε-amino group of lysine protected.

Only after acetylation step it was possible to reduce the nitro group of 3-nitrotyrosine residues to an amino group, susceptible to iTRAQ addition. Conversion of nitropeptides into their *amino* derivatives was accomplished by $Na_2S_2O_4$ treatment [46]. The peptide mixture was desalted by sep-pak C_{18} cartridge and the extent of reduction was monitored by MALDI-MS (figure II.4).

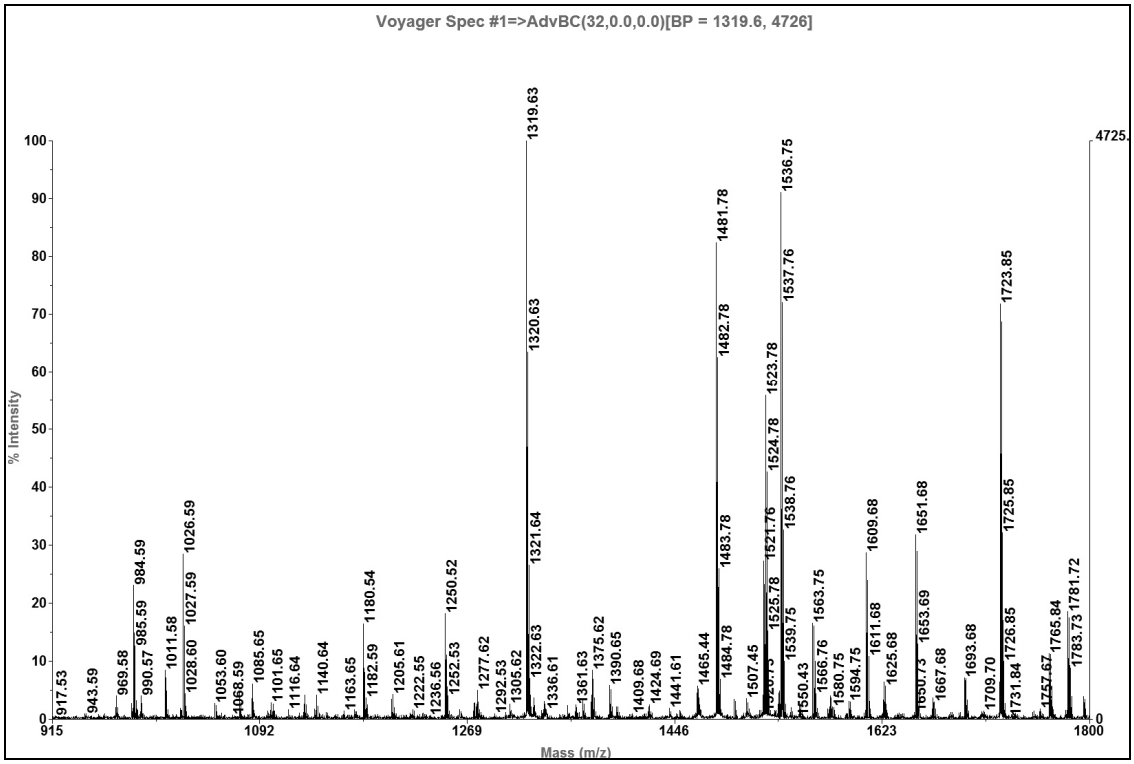


Figure II.4. *Detail of a MALDI spectrum obtained from the analysis of acetylated and reduced nitrated BSA. The spectrum shows the presence of a signal at m/z 984.59 attributable to the acetylated NH_2Tyr peptide 143-149; as a consequence of nitro group reduction, signals attributable to photodecomposition phenomena disappeared.*

In order to perform quantitative analysis, iTRAQ 114 and 117 molecules were used. Labelling reaction was carried out on the newly generated amino groups by following the protocol proposed in methods section. iTRAQ technology allows quantitative analysis of protein expression through the use of isobaric tags, enabling the quantitation of four [35] up to eight samples [47] in a single multiplexed analysis. Isobaric tagging methods are emerging as an important tool to study protein expression dynamics.

To further illustrate the ability of this labelling strategy in quantifying the relative nitration of a peptide from two different samples, iTRAQ 114 and iTRAQ 117 molecules were used to label samples containing stoichiometric concentrations of N-BSA in ratios of 1:1 and 1:4. The two samples were pooled and then submitted to LC-MS/MS analysis. We set up an experiment by using a *precursor ion scan* (PIS) for the iTRAQ reporter ions at m/z 114.1 and 117.1. In order to demonstrate that the choice of the fragment to monitor in PIS mode is un-relevant, due to the isobaric characteristic of the iTRAQ, we analysed the same sample in two separate chromatographic runs. Thus, we performed two PIS by selecting either for m/z 114.1 and m/z 117.1 fragment ions. It should be noted that PIS ion current is recorded from the fragment generated in MS^2 and no memory is retained about the intensity of the

parent ion. The chromatographic profiles are quite identical and super imposable but differing in the intensity depending on the amount of the precursor ion chosen (either 114 or 117 reporter ions), as shown in the figure II.5.

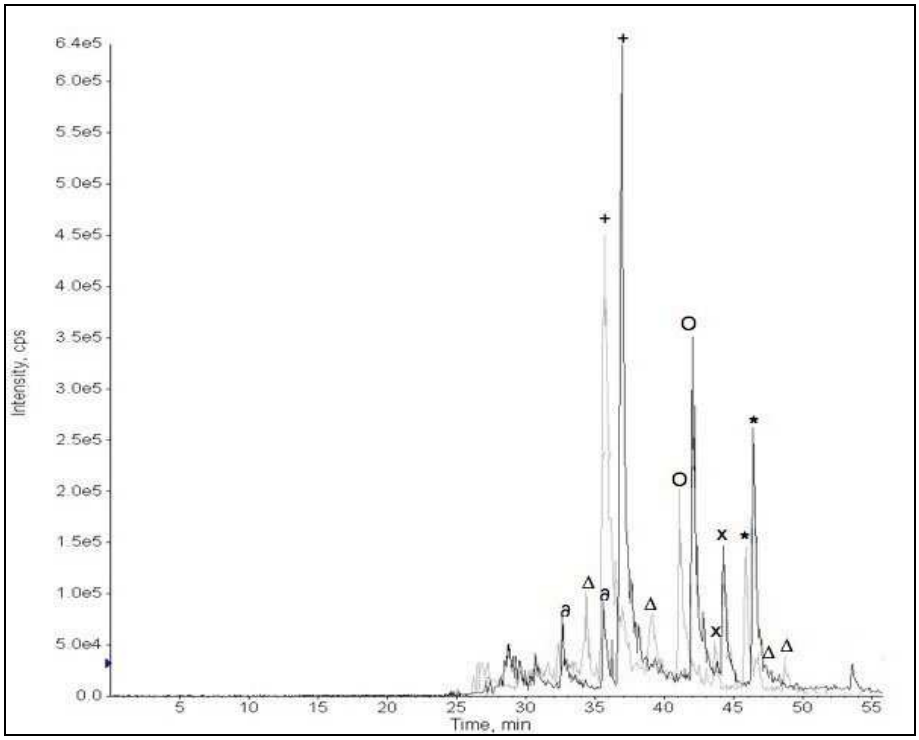


Figure II.5. *Overlapped chromatograms of nitrated BSA recorded respectively by PIS analysis of ions 114 (grey line) and 117 (bold line). The chromatogram of PIS 114 is shift back of 1 minute to lead a better comparison of signals. Symbols* ***+, O, X , *, Δ, ∂,*** *are referred respectively at BSA peptides 143-149, 403-415, 451-464, 329-341, unspecific detected in PIS114, unspecific peptide detected in PIS 117.*

To gain information about the sequence of the peptide and the site of nitration, a second stage of mass analysis is performed in which the precursor ion is selected by the first mass analyzer and fragmented. Differentially labelled peptides were indistinguishable during a single step of MS analysis because they were isobaric species due to typical characteristics of iTRAQ reagents. During MS/MS, labelled peptides produce fragmentation spectra that allow to identify the protein. Moreover, MS/MS produces fragmentation at sites modified by the iTRAQ reagent yielding reporter ions (m/z 114 and 117) that are visible as separated peaks in the low mass region of the MS/MS spectrum. The ratios of related areas of these peaks were then used to perform a relative quantitative analysis [48-49].
Figure II.6 shows the reconstructed ion chromatogram for the precursor ion scan (A). The precursor ion scan mode showed the occurrence of a number of signals, not all related to 3-nitrotyrosine-containing peptides. However, the PIS analysis led to a reduced number of signals with respect to a classical full scan MS (B) improving also the duty cycle of the method and the sensitivity toward nitro-peptides enhancing signal/noise ratio.

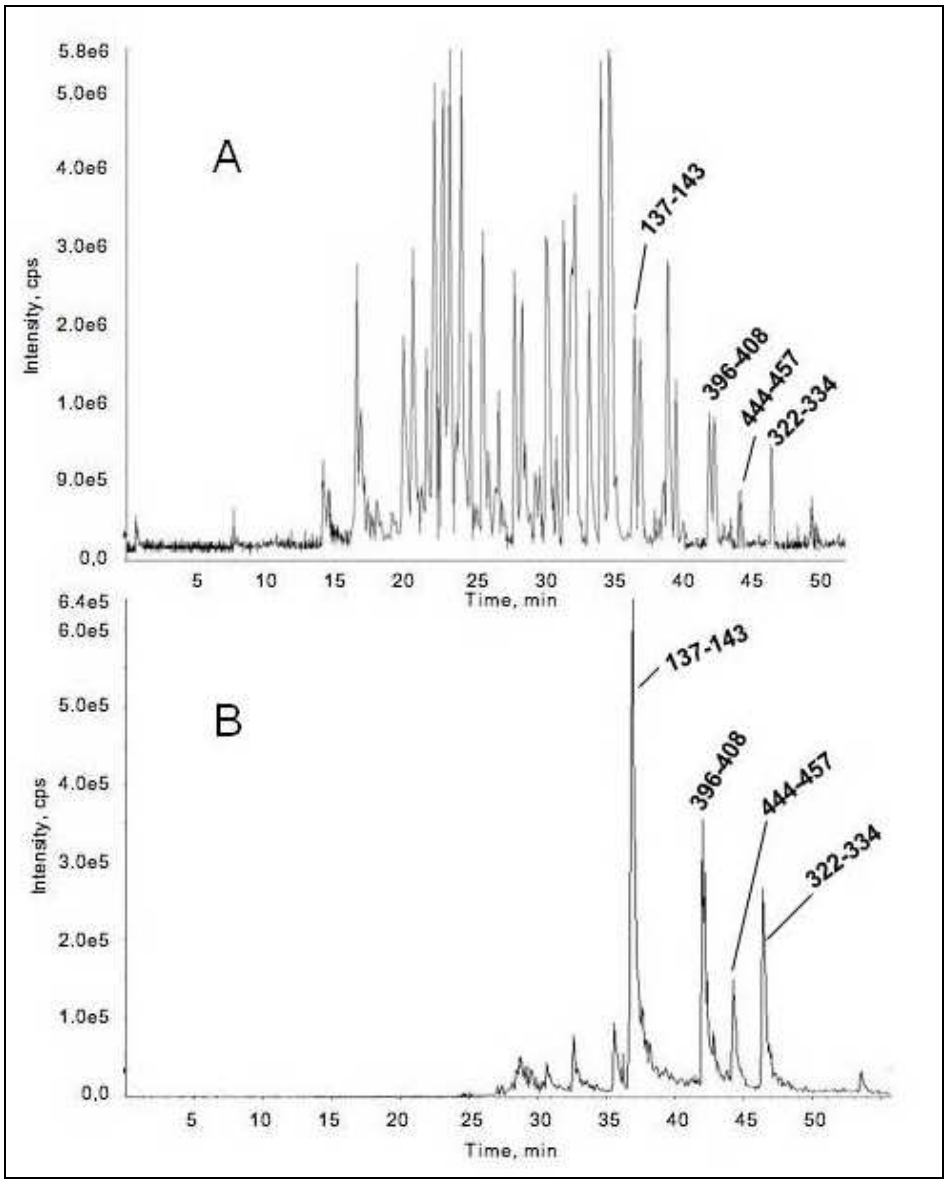


Figure II.6. Panel B *Chromatogram of nitrated BSA recorded in PIS 117.* ***Panel A*** *Chromatogram of nitrated BSA recorded in a full scan MS.*

However, others signals appeared in the chromatogram not related to nitrated peptides. This effect might be generated by the presence of different fragment ions in the region 113-114 or 116-117, many of which generated after acetylation step (table II.1). Thus I used 114 and 117 iTRAQ reagents, in order to have the worst condition of selectivity and to demonstrate the feasibility of this methodology.

Interfering ions	Mass	Peptides	Scan
(Ac)b1 Ala	114.1	(Ac)ADE(Ac)K(Ac)KFWGK(Ac)	PIS 114
(Ac)b1 Ala	114.1	(Ac)AEFVEVTK(Ac)	PIS 114
(Ac)Immonium Val	114.1	(Ac)VLASSARQR	PIS 114
(Ac)Immonium Val	114.1	(Ac)VH(Ac)KECCHGDLLECADDR	PIS 114
(Ac)Immonium Thr	116.1	(Ac)TCVADESHACCEK(Ac)	PIS 117
(Ac)Immonium Thr	116.1	(Ac)TVMENFVAFVDK(Ac)	PIS 117

Table II.1. *Un-labelled BSA peptides and their precursor detected in PIS 114 and PIS 117 modes.*

An accurate manual analysis of MS^2 spectra revealed that many false positive were selected for the presence of m/z 114 fragment whose intensity would not have effect on quantitative estimation (figure II.7).

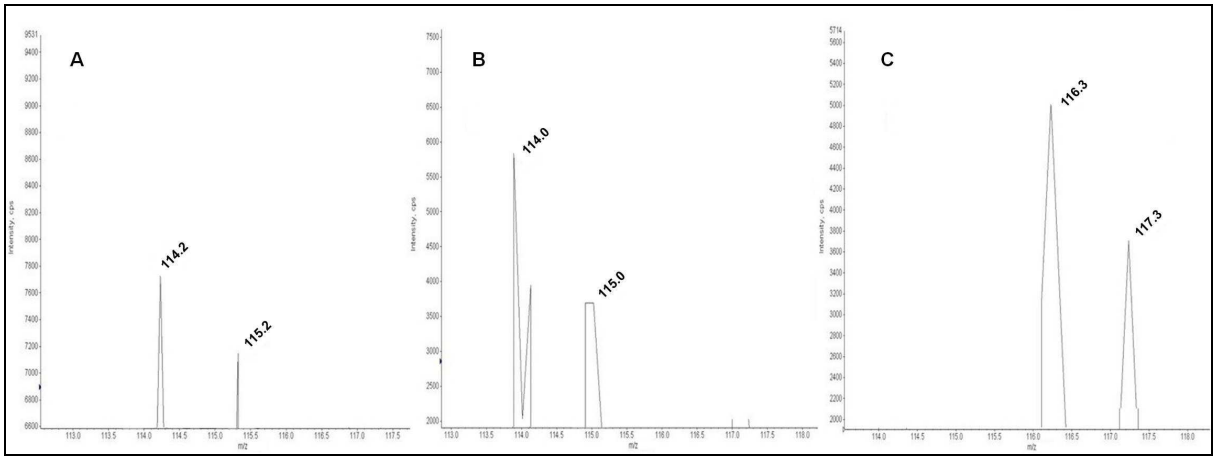


Figure II.7. *MS^2 spectra enlargement of characteristic iTRAQ m/z region relative at un-specific detected BSA peptides.* ***Panel A*** *peptide AEFVEVTK was detected in PIS114 mode because of the presence of an acetylated b1 ion of Alanine at m/z 114.2.* ***Panel B*** *peptide VLASSARQR was detected in PIS114 mode because of the presence of an acetylated immonium ion of Valine at m/z 114.0.* ***C*** *peptide TVMENFVAFVDK was detected in PIS117 mode because of the presence of an acetylated immonium ion of Threonine at m/z 116.3 or its isotope.*

To enhance the selectivity of the strategy I also tried to introduce an MS^3 event as done in previous work [32]. In particular I tried to fragment in MS^2 the ion at m/z 145 and monitor the loss of CO giving rise to the peculiar reporter ions of iTRAQ in MS^3 mode.
Preliminary experiment revealed that MS^3 spectra exhibited a very poor intensity caused by the low yield of ion fragment at m/z 145 in MS^2 mode. Thus, it was found that to enhance the selectivity of LC-MS/MS analysis good results were obtained increasing the signal threshold of PIS using the great difference in signal intensity between the iTRAQ reporter ions and others mentioned previously. However, I inferred that these LC-MS/MS conditions resulted to be a good compromise between selectivity and sensibility, fundamental parameter because of the sub-stoichiometric characteristic of protein nitration. Furthermore, the unspecific precursor selection problem is solved making a separate double PIS analysis; indeed unspecific ions detected in PIS of 114 were not detected in PIS of 117 and *vice versa* (figure II.5).

Identification of nitrated tyrosine residues was carried out by taking advantage of the flexibility of the in house Mascot software. The software was properly modified by adding in the "modification file" the value of ΔM 159 Da for iTRAQ labelled aminotyrosine residue. For Mascot analysis it was necessary to set the software supplying key information, e.g. proteolytic enzyme, fixed modifications, variable modifications, sample taxonomy, eventually missed cleavage, etc. The data obtained from MS/MS analysis in PIS mode where exploited both to identify nitration sites by using SwissProt database and to perform quantitative analysis.
Nitration sites Y^{137}, Y^{406}, Y^{457}, Y^{337} was identified by the detection and by MS/MS spectra interpretation respectively from MH_2^{2+} ions at m/z 564.8, 841.0, 963.5, 885.3.
As an example the MS/MS spectrum of modified peptide 137-143 is presented in figure II.8 A. The derivatized product ions were identified from either the corresponding b or y product ion series.
By using the Analyst software tool for calculation of peaks area it was possible to compare the areas of signals relative to reporter ions at m/z 114 and 117. For the MS/MS spectrum mentioned above was calculated a ratio of 1.1 $\pm$ 0.2.
Figure II.8 B showed the MS/MS spectrum relative to the peptide 469-482 for the mixture 1:4 molar ratios between N-BSA labelled with iTRAQ 114 and N-BSA labelled with iTRAQ 117. In this case nitration site was detectable from y ions and the area ratio for the reporter ions at m/z 117 and 114 was 4.1 $\pm$ 0.4, thus providing an accurate measurement of the relative abundance of each nitrated peptide.
It is also interesting to note in both spectra (figure II.8) the presence of an intense iTRAQ-labelled amino tyrosine immonium ion at 295.2 m/z that may be also used as indication of a nitration site presence in the peptide.

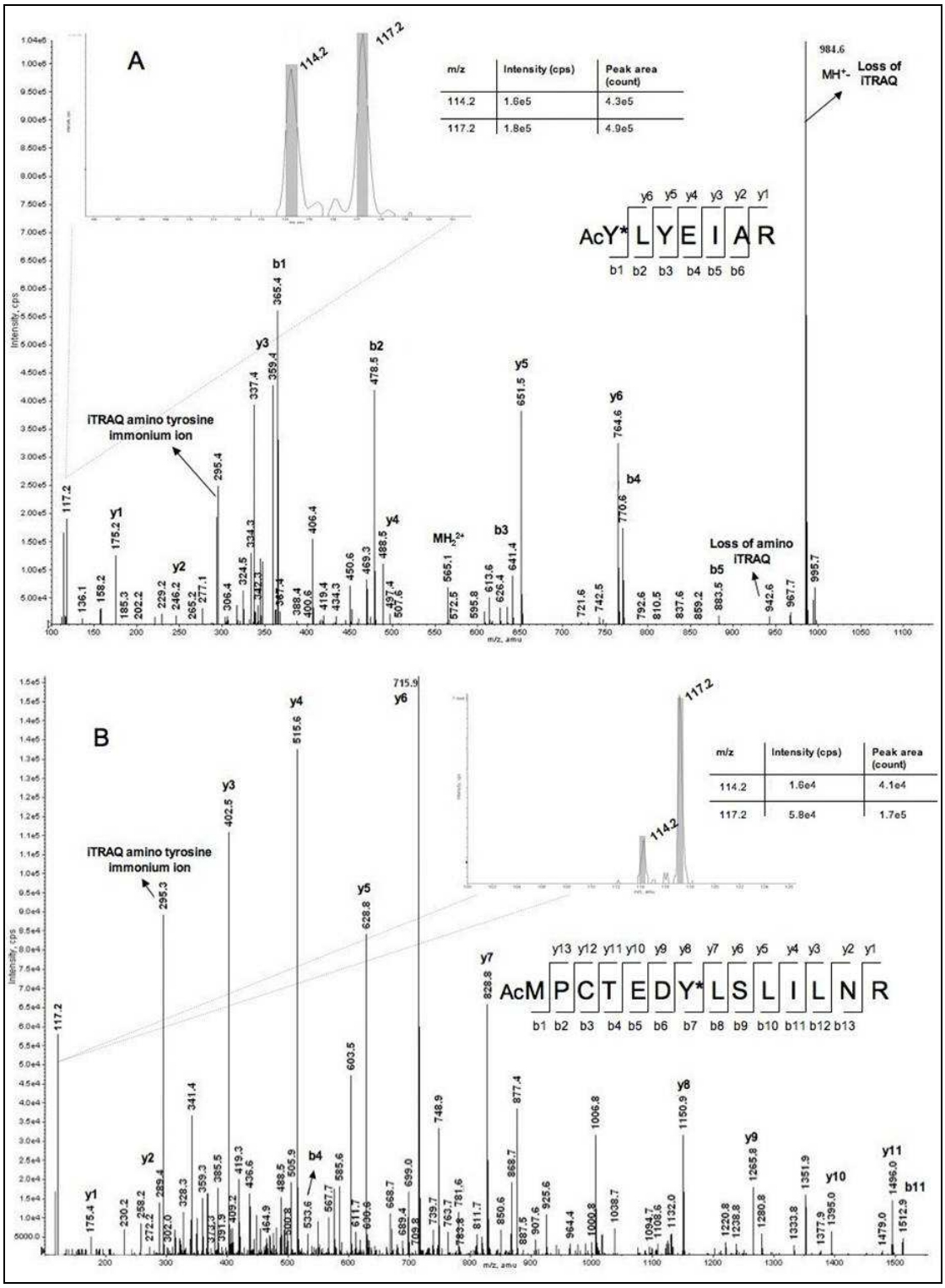


Figure II.8. Panel A *MS^2 spectrum of nitrated BSA peptide 143-149 labelled by iTRAQ reagent in 1:1 ratio.* ***Panel B*** *MS^2 spectrum of nitrated BSA peptide 451-464 labelled by iTRAQ reagent in 1:4 ratio. Spectra enlargement shows iTRAQ signals allowing quantitative estimation.*

II.3.2 Analysis of complex mixtures

In order to investigate the feasibility of applying the method to proteomics analysis, 1.5 nmol of N-BSA was added to 30 mg of an *E. coli* entire protein extract. Aliquots of mixtures of BSA and N-BSA, as control, and *E. coli* extract spiked with BSA mix were

fractionated by SDS-PAGE and submitted to Western blot analysis using anti-nitrotyrosine antibody as indicated in figure II 9.
The figure II.9 shows that the amount of N-BSA is quite impossible to differentiate from the other bacterial proteins. Moreover, in lane D is possible to appreciate that the sample used could be considered similar to a 'dirty' N-BSA *E. coli* protein extract, thus being a good candidate for nitro-proteome study.
The total protein extract was then submitted to the procedure described above. The mixture was hydrolyzed with trypsin and lysine residues were protected by acetylation. Nitrotyrosines were then reduced with dithionite, and the newly generated aminotyrosine-containing peptides were submitted to the iTRAQ labelling protocol, using 114 iTRAQ molecule.

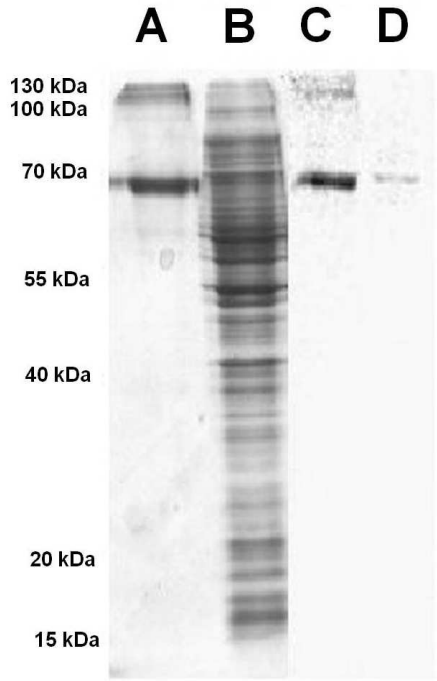


Figure II.9. *SDS-PAGE and western blot analyses of E. coli protein digest spiked with nitrated BSA (lane B) where N-BSA band was indistinguishable by the others. Lane A: Standard mix of BSA and N-BSA as control. Lanes C and D: Western blot analysis by using anti-nitrotyrosine antibody of control and E. coli samples, respectively.*

After labelling reaction, LC-MS/MS analysis of the mixture was performed as already described. In figure II.10 it was shown a three-dimensional graphic, called Counter Plot, representing LC-MS/MS parameters, time of elution, m/z and ion intensity expressed in greyscale intensity. This graphical view, obtained by an Analyst software tool, shows the duty cycle and sensitivity improvement of the strategy proposed. Figure II.10 A is referred to a full scan MS analysis of N-BSA spiked with an entire digest of *E. coli* proteins extract. As clearly shown in figure this kind of MS analysis it is not competitive for a so complex peptide mixture. Indeed, because of the large time spent by the spectrometer to analyse the more abundant peptides, lead to identify just 3 of 4 nitrated peptides present in the mixture. The high selectivity of PIS analysis proposed (figure II.10 B) reduces greatly the number of species analysed, leading the spectrometer to spend more time detecting labelled peptide, with a real gain in sensitivity.

Identification of nitrated proteins was carried out using the modified Mascot software as described above. In particular, these peptides corresponded to the same peptides previously detected in the analysis of homogeneous N-BSA. Thus, these results underscored the ability of the proposed selective labelling strategy to discriminate nitropeptides from their unmodified counterparts in a complex mixture.

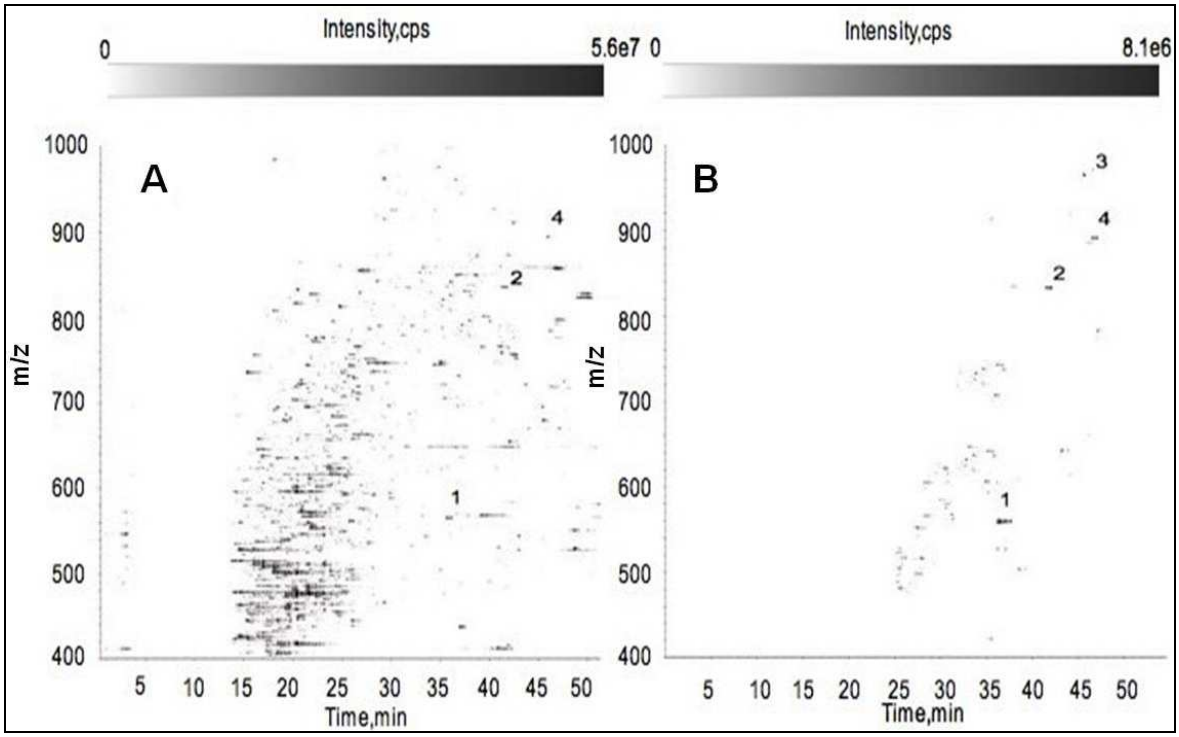


Figure II.10. *Counter plot representation of a whole E. coli protein digest spiked with nitrated BSA analyzed, by a classical full scan MS (**Panel A**) and PIS (**Panel B)** after iTRAQ selective labelling.*

The proposed strategy was finally employed to identify unknown 3-nitrotyrosine residues in a complex sample as the entire bovine milk. Bovine milk was nitrated *in vitro* with TNM [32] and the extent of nitration was monitored by SDS-PAGE and Western blot analysis as described above.
Figure II.11 (lane A) showed that the high difference of milk protein abundance in the sample reflected the high dynamic range features of the proteomic analysis. In addition the immuno-detection (lane B) showed the low extent of protein nitration thus indicating that this model system could be suitable for proteomic analysis.
The entire milk protein extract was then dissolved in denaturant buffer, and cysteine alkylation and nitro groups reduction were performed in “one pot” as described in the materials and methods section. The protein sample was then desalted and digested with trypsin. The resulting peptide mixture was acetylated, divided in ratio 1:1 and 1:4 then selectively labelled with iTRAQ reagents and analysed by nanoLC-MS/MS in PIS mode either for the ion 114 or for the ion 117.

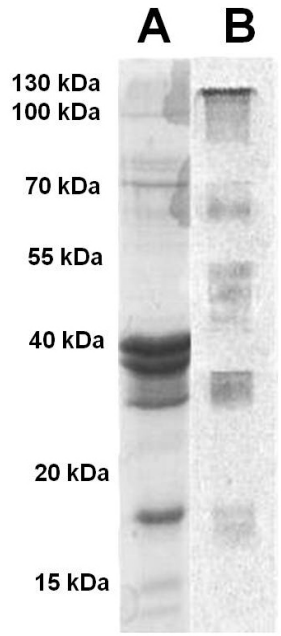


Figure II.11. *SDS-PAGE (lane A) and Western blot analysis by using anti-nitrotyrosine antibody (lane B) of in vitro nitrated milk proteins.*

Identification of nitrated milk proteins was carried out using the modified Mascot software as described. The common results obtained from a triplicate analysis of the two PIS mode are summarised in the table II.2.

Observed m/z	Score	Sequence	Ratio 114:117	Protein, accession no.
740.2	59	(Ac)VLVDTY*K(Ac)K(Ac)	4.1 ± 0.1 1.2 ± 0.2	β-Lactoglobulin, P02754
796.5	49	(Ac)DMPIQAFLLY*QEPVLGPVR	3.8 ± 0.4 0.9 ± 0.2	β-Casein, P02666
573.5	44	(Ac)TVY*QHQK(Ac)	4.3 ± 0.2 1.2 ± 0.3	α-S1-Casein, P02663
568.8	42	(Ac)Y*LYEIAR	3.9 ± 0.1 0.8 ± 0.3	Albumin bovine, P02769
426.8	36	(Ac)ADLIAY*LK(Ac)K(Ac)	4.1 ± 0.4 1.3 ± 0.2	Cytochrome c, P62849
716.3	34	(Ac)STRTVSSSSY*R	4.1 ± 0.4 1.2 ± 0.4	Vimentin, P48616
617.4	30	(Ac)AY*PTPARSK(Ac)	4.2 ± 0.7 1.4 ± 0.3	Glycoprotein hormones a chain, P01217
565.9	29	(Ac)HFHLY*GR	5.2 ± 1.6 1.8 ± 0.9	Integrin α-L, P61625
559.8	29	(Ac)LELY*LPK(Ac)	7.1 ± 3.4 1.8 ± 1.2	Plasma serine protease inhibitor, Q9N212

Table II.2. *Nitrated milk Proteins detected by the selective iTRAQ labelling strategy coupled with a selective PIS-MS analysis.*

As indicated in the table, besides the high-abundant milk proteins (α-casein and β-lactoglobulin), this procedure was also able to assess the nitration sites occurring in low-abundant proteins, like Vimentin and Integrin, thus demonstrating the feasibility of this strategy for the identification of protein nitration in proteomics. Concerning the quantifications we realized the quantitative estimation of 7/9 nitrated peptides. Indeed, if from a side the less abundant proteins spectra lead to unequivocally identify the sequence, the quality of signals for the ions 114 and 117 were too low to make an accurate quantitative measure.
In figure II.12 it is shown the MS/MS spectrum relative to the nitrated α-casein, it is interesting noting that the presence of an acetylated threonine in N-terminal position has not effect on the quantitative estimation.

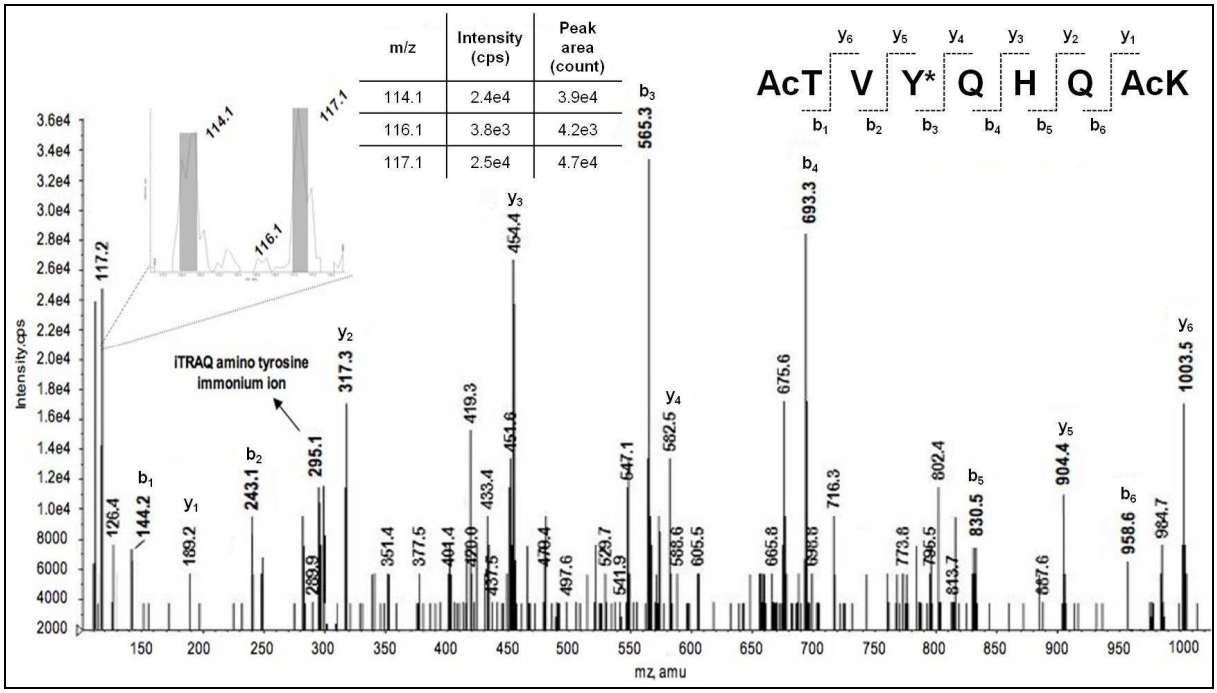


Figure II.12. *MS^2 spectrum relative at α-Casein nitrated peptide TVYQHQK detected in an in vitro nitrated milk proteins extract. In the spectrum enlargement of iTRAQ mass region it is visible the characteristic signals of the reporter and also a signal at 116.1 m/z relative at Acetylated Threonine immonium ion. The presence of this fragment has not effect on the quantitative estimation.*

II.4 Conclusions

The iTRAQ approach has been successfully applied to a variety of prokaryotic and eukaryotic samples including *Escherichia coli*, yeast, human saliva, human fibroblasts and mammary epithelial cells to identify and quantify the proteins in these samples [50-51]. Zhang *et al*. [52] combined the iTRAQ labelling approach with immunoprecipitation to quantify tyrosine phosphorylated peptides in epithelial cells. However, iTRAQ has been applied to identify and quantify protein level in shotgun proteomics approach [53-54]. The work presented in this chapter represents the first innovative application of iTRAQ labelling strategy to the analysis of post-translational modifications. In particular it was addressed the analysis of protein nitration, thus providing a powerful analytical methods to identify nitrated protein, localise tyrosine nitration sites and quantitate the extent of protein nitration in a single experiment.

Here the strategy proposed, taking in account the availability of QqLIT instruments to select specific labelled peptides giving rise to diagnostic MS^2 product ions, resulted to be of interest in the selective detection of protein nitration. Moreover, because of its operational simplicity, avoiding long-lasting and time-consuming fractionation procedures, this new strategy seems to be well suited for large-scale proteomic profiling of nitration sites.
The strategy reported here was set up with a 4000QTrap mass spectrometer; however, this approach may be realised also using others mass spectrometer optics, with opportune modifications. For example, Niggeweg et al. [55] showed the possibility to perform PIS compatible with chromatographic separations by a QTOF. Moreover, it would be very interesting to evaluate possible improvements, in terms of selectivity and quantitative estimation, by measuring with high resolution and accuracy (< 10 ppm) iTRAQ reporter ions. This may be accomplished by a Thermo LIT/Orbitrap performing MS/MS experiments in higher energy CID (HCD) [56]. The optic and the electronics of this instrument lead to perform analysis in an operation mode called "Parent Ion Mapping" that may be considered as reconstructed PIS. However, our feeling is that, because of the higher duty cycle of a triple quadrupole-based PIS, eventual improvement would have real effectiveness in an offline LC-MALDI-MS/MS analysis more than a classical LC-ESI-MS/MS.

II.5 References

1. Lu, Y.; Bottari, P.; Aebersold, R.; Turecek, F.; Gelb, M.H. *Methods Mol Biol.* **2007**, *359*, 159-176.
2. Foettinger, A.; Leitner, A.; Lindner, W.; *J Proteome Res.* **2007**, *6*, 3827-3834.
3. Amini, A.; Chakraborty, A.; Regnier, F.E. *J Chromatogr B Analyt Technol Biomed Life Sci*. **2002**, *772*, 35-44.
4. Temple, A.; Yen, T.Y.; Gronert, S. *J Am Soc Mass Spectrom*. **2006,** *17,* 1172-1180.
5. Camerini, S.; Polci, M.L.; Restuccia, U.; Usuelli, V.; Malgaroli, A.; Bachi, A. *J Proteome Res.* **2007**, *6*, 3224-3231.
6. Jalili, P.R.; Sharma, D.; Ball, H.L. *J Am Soc Mass Spectrom*. **2007**, *18*, 1007-1017.
7. Mirzaei, H.; Regnier, F. *Journal of Chromatography B* **2005**, *817*, 23-34.
8. Zhang, Q.; Qian, W.J.; Knyushko, T.V.; Clauss, T.R.; Purvine, S.O.; Moore, R.J.; Sacksteder, C.A.; Chin, M.H.; Smith, D.J.; Camp, D.G.; Bigelow, D.J.; Smith, R.D. *J Proteome Res.* **2007**, *6*, 2257-2268.
9. Zhang, H.; Li, X.J.; Martin, D.B.; Aebersold, R. *Nat Biotechnol*. **2003**, *21*, 660-666.
10. Ischiropoulos, H. *Arch. Biochem. Biophy* **1998***, 356,* 1-11.
11. Eiserich, J.P.; Patel, R.P.; O'Donnell, V.B. *Mol Aspects Med.* **1998**, *19*, 221-357.
12. Shi, W.Q.; Cai, H.; Xu, D.D.; Su, X.Y.; Lei, P.; Zhao, Y.F.; Li, Y.M. *Regul Pept*. **2007**, *144*, 1-5.
13. Lin, H.L.; Myshkin, E.; Waskell, L.; Hollenberg, P.F*. Chem Res Toxicol.* **2007**, *20*, 1612-1622.

14. Lopez, C.J.; Qayyum, I.; Mishra, O.P.; Delivoria-Papadopoulos M. *Neurosci Lett.* **2005**, *386*, 78-81.
15. Lennon, C.W.; Cox, H.D.; Hennelly, S.P.; Chelmo, S.J.; McGuirl, M.A. *Biochemistry* **2007**, *46*, 4850-4860.
16. Zhan, X.; Desiderio, D.M., *Biochem Biophys Res Commun.* **2004**, *325*, 1180–1186.
17. Kanski, J.; Behring, A.; Pelling, J.; Schöneich C*. Am J Physiol Heart Circ Physiol.* **2005**, *288*, H371–H381.
18. Tedeschi, G.; Cappelletti, G.; Negri, A.; Pagliato, L.; Maggioni, M.G.; Maci, R.; Ronchi, S. *Proteomics.* **2005**, *5*, 2422-2432.
19. Söderling, A.S.; Ryberg, H.; Gabrielsson, A.; Lärstad, M.; Torén, K.; Niari, S.; Caidahl, K. *J Mass Spectrom.* **2003**, *38*, 1187-1196.
20. Krueger, R.C. *Anal Biochem.* **2004**, *325*, 52-61.
21. Sacksteder, C.A.; Qian, W.J.; Knyushko, T.V.; Wang, H.; Chin, M.H.; Lacan, G.; Melega, W.P.; Camp, D.G.; Smith, R.D.; Smith, D.J.; Squier, T.C.; Bigelow, D.J. *Biochemistry.* **2006**, *45* (26), 8009-8022.
22. Thomson, L.; Christie, J.; Vadseth, C.; Lanken, P.N.; Fu, X.; Hazen, S.L.; Ischiropoulos, H.; *Am J Respir Cell Mol Biol*. **2007**, *6* (2), 152-7.
23. Zhan, X.; Desiderio, D.M. *Anal Biochem*. **2006**, *354* (2), 279-89.
24. Nikov, G.; Bhat, V.; Wishnok, J.S.; Tannenbaum S.R. *Anal Biochem*. **2003**, *320*, 214-222.
25. Zhang, Q.; Qian, W.J.; Knyushko, T.V.; Clauss, T.R.; Purvine, S.O.; Moore, R.J.; Sacksteder, C.A.; Chin, M.H.; Smith, D.J.; Camp, D.G.; Bigelow, D.J.; Smith, R.D. *J Proteome Res*. **2007**, *6*, 2257-2268.
26. Sarver, A.; Scheffler, N.K.; Shetlar, M. D.; Gibson, B. W. *J Am Soc Mass Spectrom*. **2001**, *12* (4), 439–48.
27. Petersson, A. S.; Steen, H.; Kalume, D. E.; Caidahl, K.; Roepstorff,P. *J Mass Spectrom*. **2001**, *36* (6), 616–25.
28. Petre, B. A.; Youhnovski, N.; Lukkari, J.; Weber, R.; Przybylski, *M. Eur. J. Mass Spectrom*. **2005**, *11* (5), 513–8.
29. Turko, I. V.; Murad, F. *Methods Enzymol*. **2005**, *396*, 266–75.
30. Zhan, X.; Desiderio, D. M. *Int J Mass Spectrom*. **2007**, *259* (1-3), 96–104.
31. Radi, R. *Proc Natl Acad Sci U.S.A*. **2004**, *101* (12), 4003–4008.
32. Amoresano, A.; Chiappetta, G.; Pucci, P.; D'Ischia, M.; Marino G. *Anal Chem.* **2007**, *79*, 2109-2117.
33. Chen, X.; Sun, L.; Yu, Y.; Xue, Y.;Yang, P. *Expert Rev Proteomics* **2007**, *4*, 25-37.
34. Gygi, S.P.; Rist, B.;, Gerber, S.A.; Turecek, F.; Gelb, M.H.; Aebersold, R. *Nat Biotechnol.* **1999**, *17*, 994-999.
35. Ross, P.L.; Huang, Y.N.; Marchese, J.N.; 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, D.J. *Mol Cell Proteomics*. **2004**, *3*, 1154-1169.

36. Ow, S.Y.; Cardona, T.; Taton, A.; Magnuson, A.; Lindblad, P.; Stensjö, K.; Wright, P.C. *J Proteome Res.* **2008**, *7*, 1615-1628.
37. Pettersson, A. S.; Steen, H.; Kalume, D. E.; Caidahl, K.; Roepstorff, P. *J. Mass Spectrom.* **2001**, *36*, 616-625.
38. Sokolovsky, M.; Riordan, J. F.; Vallee, B. L. *Biochem Biophys Res Commun.* **1967**, *27*, 20-25.
39. Staudenmann, W.; James, P. *Proteome research: Mass Spectrometry*. **2001,** 160-161.
40. Jankolovits, J.; *Journal of Young Investigators* **2008**, *15*.
41. Jencks, W.J. *Acc Chem Res.* **1976,** *9,* 425-432.
42. Geng, M; Ji, J.; Regnier, F. E. *J Chromatogr A.* **2000**, *870*, 295-313.
43. Brancia, F. L.; Oliver, S. G.; Gaskell, S. J. *Rapid Commun Mass Spectrom.* **2000**, *14*, 2070-2073.
44. Zhang, X.; Jin, Q. K.; Carr, S. A.; Annan. R. S. *Rapid Commun Mass Spectrom.* **2002**, *16*, 2325-2332.
45. Zappacosta, F.; Annan, R. S. *Anal Chem.* **2004**, *76*, 6618-6627.
46. Sokolovsky, M.; Riordan, J. F.; Vallee, B. L. *Biochemistry.* **1966**, *5,* 3582-3589.
47. Choe, L.; D'Ascenzo, M.; Relkin, N. R.; Pappin, D.; Ross, P.; Williamson, B.; Guertin, S.; Pribil, P.; Lee, K.H. *Proteomics.* **2007**, *20,* 3651-3660.
48. Skalnikova, H.; Rehulka, P.; Chmelik, J.; Martinkova, J.; Zilvarova, M.; Gadher, S. J.; Kovarova, H. *Anal Bioanal Chem.* **2007,** *5*,1639-1645.
49. Gafken, P. R.; Lampe, P. D. *Cell Commun Adhes*. **2006**, *5-6*, 249-262.
50. Hardt, M.; Witkowska, H.E.; Webb, S.; Thomas, L.R.; Dixon, S.E.; Hall, S.C.; Fisher, S.J.; *Anal Chem.,* **2005**, *77*, 4947-4954.
51. Cong, Y.S.; Fan, E.; Wang, E. *Mech Ageing Dev.* **2006**, *127*, 332-343.
52. Zhang, Y.; Wolf-Yadlin, A.; Ross, P. L.; Pappin, D. J.; Rush. J.; Lauffenburger, D. A.; White, F. M. *Mol Cell Proteomics* **2005,** *9*, 1240-1250.
53. Aggarwal, K.; Choe, L.H.; Lee, K.H.; *Brief Funct Genomic Proteomic* **2006**, *5*, 112-120.
54. Jorrín, J.V.; Maldonado, A.M.; Castillejo, M.A. *Proteomics.* **2007**, *7*, 2947-2962.
55. Niggeweg, R.; Köcher, T.; Gentzel, M.; Buscaino, A. *Proteomics*. **2006**, *6*, 41-53.
56. Bantscheff, M.; Boesche, M.; Eberhard, D.; Matthieson, T. *Mol Cell Proteomics*. **2008**, *7*, 1702–1713.

III. DANSYL LABELLING AND BIDIMENSIONAL MASS SPECTROMETRY TO INVESTIGATE PROTEIN CARBONYLATION

III.1 Introduction

In this chapter it is reported a new methodology for selective analysis of carbonylated residues, and in particular of carbonylated lysine residues. As for protein nitration analysis, the strategy proposed may be classified as belonging to the class of RIGhT methodologies (section I.2.4). The approach proposed is based on the labelling of target residues with reagents capable to generate reporter ions in MS^2/MS^3 experiments. The advantages of this analytical strategy are reported in the previous chapter (section II.1). The results discussed in the previous chapter showed that by using a very selective MS analysis it is possible to improve both the dynamic range of the mass spectrometer and the signal/noise ratio. In operative terms, the use of a high selective MS replaces the off-line chromatographic steps (affinity and/or bidimensional chromatography) aimed at isolating the labelled peptides, in order to avoid the time consuming analysis and the tremendous decrease in sensitivity. The strategy proposed here can be considered as a third generation proteomics approach. The method proposed was proven to lead the precise localisation of carbonylated residues in proteins.

III.1.1 Proteomics analysis of carbonylated proteins

Carbonylation of proteins may be the most widely type of damage used to infer oxidative stress [1-3], mainly because carbonyl modifications can be induced by a wide variety of ROS as well as by by-products of lipid oxidation (section I.3.2.2).
The increase in carbonylated proteins during cellular ageing or in response to oxidative stress is not a casual event; in fact some proteins are more susceptible to ROS than others; moreover most of these proteins are specie specific [4-9]. Thus the definition of the entire set of target proteins of oxidative stress is a hard challenge. For these reasons analysis of protein carbonylation is not easy and several methods have been developed to determine the content of carbonyl groups in proteins [10-15]. Protein carbonyl groups have no distinguishing spectrophotometric absorbance/fluorescence properties, so they cannot be directly determined. For this reason, detection and quantification of carbonyl groups of proteins require the use of specific chemical probes. A simple and accurate method to quantify protein oxidation uses 2,4-dinitrophenylhydrazine, which specifically reacts with carbonyls to form 2,4-dinitrophenylhydrazone [16-17]; the amount of hydrazone is quantified by using spectrophotometric technique. An important caveat to be considered when using this method is that proteins containing chromophores (i.e. heme-containing proteins) have absorbance wavelengths similar to DNPH and may interfere with its measurement [18], leading to inaccurate estimation of protein carbonyls. The DNPH assay has also been used in combination with SDS-PAGE and Western blot analysis using commercially available anti-DNPH antibodies, thus enhancing the sensitivity of the method [19-22]. Initial studies adopting SDS-PAGE approach led to the observation that not all proteins in a given proteome were subject to equivalent oxidative chemical modification, supporting the current view that protein oxidation is a selective process [23-24]. More recent studies have successfully analyzed DNPH-modified samples using 2D-GE in a variety of experimental systems [25-29]. Using fluorescent

secondary antibodies, protein carbonyl groups derivatized with DNPH can also be visualized through *in situ* immunofluorescent imaging [30-34]. Using this strategy it is possible to go further insights into the mechanisms of cellular oxidative stress during physiological and pathological conditions [35-36]. An excellent example of *in situ* immunofluorescent imaging of carbonylated proteins is represented by the work of Desnues *et al.* [32], in which the level of protein carbonylation in whole *Escherichia coli* cells was shown to be greatly elevated by H_2O_2 treatment.

As an alternative to DNPH treatment, fluorescent methods have been developed; in particular two visible-wavelength fluorescence probes have been developed based on the use of fluorescein-5-thiosemicarbazide (FTC) [37-39] and Alexa 488 fluorescence hydroxylamine (FHA, Invitrogen) [40]. Both probes are commercially available and have been successfully used in 2D-GE based proteomics. Application of these probes to quantify carbonyl content in protein mixtures has been limited, presumably because of the fact that it is difficult to completely remove the excess probes. Moreover the use of 2D-GE based analytical strategies is limited by the drawbacks of this analytical tool. A unique probe for this kind of analysis is APTB that is not fluorescent by itself but becomes fluorescent upon reaction with carbonyl groups [41]. Using these probe protein carbonyl content could be read directly from protein mixtures with a fluorometer, because the unconjugated excess APTB is non-fluorescent. The use of fluorescence probes offers at least two advantages when used in combination with gel-based proteomics. First of all, there are no Western blot experiments to be carried out, so the whole process can be completed in a shorter time. Secondly, the same gel can be used for both protein staining and protein carbonyl imaging [37-38]; this reduces errors associated with gel spot identification and excision for subsequent mass spectrometric analysis, while using DNPH gels are used for total protein staining and Western blot membranes are used for imaging of the carbonylated proteins [25-26].

A radioactive method has been developed that uses tritiated sodium borohydride ($NaBH_4$) to reduce carbonyl groups and their Schiff bases; in this case tritium is incorporated selectively into the carbonylated proteins [42] and the radioactivity can be measured to quantitate the carbonyl content of proteins. This strategy is also compatible with SDS-PAGE separation of labelled proteins; in this case bands of interest can be excised and identified [43]. This methodology can be used when target proteins have absorbance near 360 nm that would otherwise interfere with DNPH measurement, but has some drawbacks such as the use of radioactive reagents, with risks for safety, and non specific incorporation of tritium into proteins [42].

Recently proposed methods for detection of carbonyl groups use biotin for carbonyls labelling; biotin probes take advantage of the affinity binding between biotin and avidin, among the strongest known non-covalent interactions [44]. The most used biotin probe for the analysis of protein carbonyls is biotin-hydrazide [4]. Proteins labelled with this probe can be either affinity captured using avidin beads [13] or analysed using 2D-GE approaches. In this case the advantage is that no secondary antibodies are required when biotin-containing probes are used for Western blot detection of carbonyls. Using this probe the initial reaction between a carbonyl group and the hydrazide forms a hydrazone moiety that is unstable and must be reduced to a more stable hydrazine group [4]. This strategy has been used in combination with affinity chromatography step and mass spectrometric analysis to identify specific sites of carbonylation in proteins [45-49]. An excellent application of this probe is the profiling of carbonylated proteins in human plasma [13]. In this study, the use of biotin-hydrazide

allowed accurate estimation of total carbonylated proteins in plasma of healthy humans (0.2%).
All the strategies described could be used in combination with mass spectrometric identification of carbonylated proteins. As for gel-based approach, bands (1D-GE) or spots (2D-GE), whether detected by chemiluminescence, radioactivity, or fluorescence, can be analysed by LC–MS/MS techniques following *in situ* trypsin digestion. The major drawbacks of this method are linked to the limits of gel-based proteomics (see section I.2.2); moreover this approach will only identify proteins that are putatively carbonylated [26,50] and it suffers from the difficulty to pinpoint the modified residues unless peptides laying carbonylated residues are sequenced. Unfortunately this is usually the case for complex samples that do not receive prior selective enrichment of modified species, since PTMs are sub-stoichiometric modifications. Therefore to selectively identify the proteins that undergo carbonylation, a shotgun proteomic approach may be used [51-52]. This means that protein carbonyl groups have to be derivatized with biotin-containing probes, then digested by trypsin and enriched by affinity capture of the carbonylated/tryptic peptides, using avidin beads, prior to mass spectrometric analysis [13,51,53-54]. The use of biotin probes facilitates isolation and enrichment of the carbonylated peptides and, at the same time, tags the carbonylated sites that can be distinguished by mass spectrometry. However biotin tags can diminish the ionization efficiency and can alter the fragmentation pattern of modified peptide causing a loss in signal intensities and in the quality of MS/MS spectra.
The research group in which i worked for this PhD thesis have recently reported innovative RIGhT (Reporter Ion Generating Tag) approaches, which involve dansyl labelling of nitration and phosphorylation sites and rely on the enormous potential of MS^n analysis [55-56]. Taking advantage of the experience already made, this chapter reports a novel RIGhT methodology for selective labelling of carbonyl groups in proteins using dansylhydrazide coupled with selective analysis by *precursor ion scan* mode and MS^3. This strategy is certainly rapid and easy to use but suffers from the impossibility to perform quantitative analysis of samples of interest because of the absence of heavy form of dansyl-hydrazide molecule. For this reason in the last part of this chapter it is reported a preliminary approach that exploits the potential of iTRAQ reagents to perform selective and quantitative analysis of carbonyl groups of proteins.

III.1.2 Dansyl derivatives as RIGhT reagent

By using reagents as iTRAQ or Dansyl (DNS) it is possible to design analytical strategies defined as RIGhT approaches [56]; these approaches are based on the separation of analytes by using the molecular structure features of selectively labeled peptides and the fragmentation characteristics introduced by the reagent chosen, rather than using chromatographic or electrophoretic fractionation. RIGhT reagents (as described in section II.1.2) are characterized by having a reactive group toward the residues of interest, a linker group that is loss as neutral moiety in MS^2 experiment and a reporter group that generates stable fragment ions selectively detectable in MS^2 experiment.
For the development of a RIGhT strategy to selectively identify carbonylated residues of proteins, *dansyl-hydrazide* (DNSH) has been chosen as molecular probe. DNSH is a specific carbonyl reactive molecule with all the structural features of a RIGhT reagent (figure III.1).

The choice of this reagent was due to the well known fragmentation features of dansyl-derivatives in gas phase [56]. My tutor's research group has a relevant experience in the usage of dansyl-derivatives in the years before the commercial availability of mass spectrometers. Since the introduction of these instruments, at the end of the 60th, MS characteristics of dansyl derivatives have been extensively studied; in the 1968 Marino *et al*. showed fragmentation features of dansyl-derivatives during electronic impact fragmentation events [57].

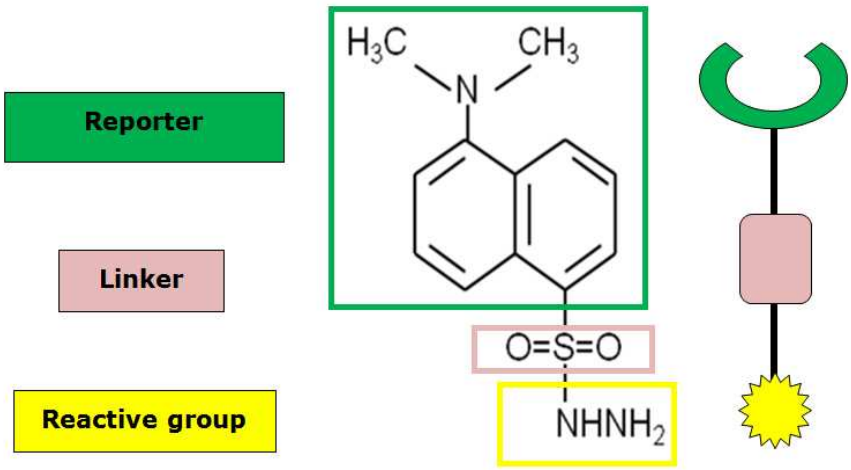


Figure III.1. *Dansyl-hydrazide has the peculiar structure of RIGhT reagents.*

During MS^2 fragmentation, dansyl peptides give rise to two specific fragment ions at 170 m/z and 234 m/z; in addition the last one may be further fragmented in a diagnostic MS^3 transition 234-170 m/z (figure III.2).
These features were exploited by using characteristic scan modes of an Applied Biosystems 4000QTrap mass spectrometer. By coupling this instrumentation and dansyl-derivatives it was possible to set up a high throughput LC-MS/MS method based on *Precursor Ion Scan* (PIS) and MS^3 experiments. PIS consists in: i) a Q_1 scan across the full mass range; ii) ions separated are fragmented in the collision cell (q_2); iii) Q_3 is set to transmit only ions with the specific mass of the diagnostic fragment m/z 170. Selected precursor ions are then submitted to combined MS^2/MS^3 experiments, to specifically detect only those ions giving rise to the transition 234-170 in MS^3 mode,

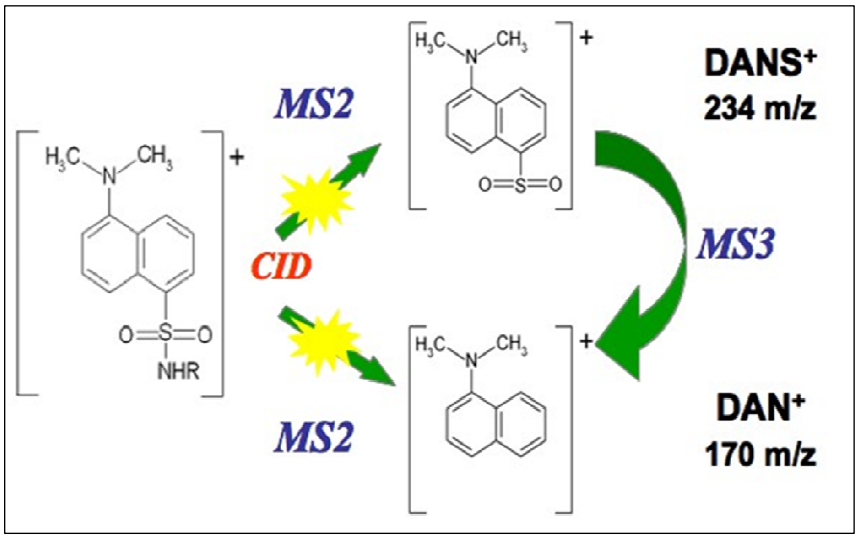


Figure III.2. *Peculiar dansyl-peptides CID fragmentations.*

thus leading to the selective detection of dansyl-peptides. This MS based selection of labelled peptides substitutes other chromatographic steps so that the mass spectrometer can be considered as the second dimension of the RP-LC separation.

III.2 Materials and methods

Chemicals

Dithiothreitol (DTT), ethylendiaminetetraacetic acid (EDTA), trypsin, urea, α-cyano-4-hydroxycinnamic acid, sodium hypochlorite solution, bovine ribonuclease A (RNase A), lysozime and hydrazine solution were from Sigma (St. Louis, MO). 5-N,N-(dimethylamino)naphthalene-1-sulfono-hydrazide (DNSH), ammonium hydrogen carbonate (AMBIC) and iodoacetamide (IAM) were purchased from Fluka. Acetonitrile (ACN) was from Romil. Trifluoroacetic acid HPLC grade was purchased from Carlo Erba. All solvents were of the highest purity available from Baker. All other reagents and proteins were of the highest purity available from Sigma.

In vitro carbonylation of RNase A

RNase A was dissolved in a 5 mM NaClO solution to a final concentration of 0.5 mg/ml. The reaction was carried out at 37 ℃ for 15 min. Carbonylated RNase A was immediately desalted by size exclusion chromatography on a Sephadex G-25M column (GE Healthcare). Protein elution was monitored at 220 and 280 nm. Fraction containing protein was collected, lyophilized and stored at -20 ℃.

Disulfide bridges reduction, Cys alkylation and tryptic digestion

Carbonylated RNase A was dissolved in denaturant buffer (urea 6M, Tris 300 mM pH 8.0, EDTA 10mM) and disulfide bridges were reduced with DTT (10-fold molar excess on the Cys residues) at 37 ℃ for 2 h and th en alkylated by adding IAM (5-fold molar excess on thiol residues) at room temperature for 30 min in the dark. Protein samples were desalted by size exclusion chromatography on a Sephadex G-25M column (GE Healthcare). Fraction containing protein was lyophilized and the dissolved in 50 mM AMBIC buffer pH 8.0. Trypsin digestion was performed using an enzyme/substrate ratio of 1/50 w/w at 37 ℃ for 16 h.

Labelling of carbonyl groups with DNSH

Sample was lyophilized and then dissolved in 10 μl of H_2O. To this solution it was added 100 μl of a 6.5 % of TFA solution in ACN and an equal volume of a 5 mM DNSH solution in ACN. Solutions were allowed to react at 37 ℃ for 8 h.

Bacterial strains, growth conditions and proteins extract preparation

Escherichia coli K12 strain was grown in aerobic conditions at 37 °C in LB medium. After 16 h, bacteria were harvested by centrifugation and resuspended in Buffer Z (25 mM HEPES pH 7.6, 50 mM KCl, 12.5 mM $MgCl_2$, 1 mM DTT, 20% glycerol, 0.1% triton) containing 1 mM PMSF. Cells were disrupted by sonication. The suspension was centrifuged at 9000 x g for 30 min at 47 ℃. Af ter centrifugation the protein concentration of the extract was determined by Bradford assay.

iTRAQ-hydrazide synthesis

iTRAQ-hydrazide was obtained by reaction of iTRAQ (Applied Biosystems, Framingham, MA) (1 nmol/μl dissolved in ethanol) with hydrazine (50 nmol/μl in ethanol) (molar ratio 2:1). Reaction was carried out at room temperature for 6 h at

room. Product was purified by RP-HPLC using an Agilent Zorbax C18 column (4,6mm X 150mm) (Palo Alto, California) using a 10% to 65% linear gradient in 20 min from water to acetonitrile (ACN) and verified by MALDI-MS analysis.

Labelling of carbonyl groups with iTRAQ-hydrazide

Lyophilized peptide sample was dissolved in 10 µl of H_2O. To this solution it was added 100 µl of a 6.5 % of TFA solution in ACN and an equal volume of a 5 mM iTRAQ-hydrazide solution in ACN. Solutions were allowed to react at 37 ℃ for 18 h. After reaction sample was concentrated and diluted in HCOOH 0.1% for mass spectrometric analysis.

Mass Spectrometry

MALDI-MS experiments were performed on a Voyager-DE STR MALDITOF mass spectrometer (Applied Biosystems, Framingham, MA) equipped with a nitrogen laser (337nm). 1 µl of the total mixture was mixed (1/1, v/v) with a 10 mg/ml solution of α-cyano-4-hydroxycinnamic acid in ACN/50 mM citrate buffer, 70/30 v/v.

MALDI-MS/MS analysis of protein mixtures labelled with iTRAQ-hydrazide, was performed on a MALDI TOFTOF 4800 Plus mass spectrometer (Applied Biosystems, Framingham, MA). 0.5 µl of the total mixture was mixed (1/1, v/v) with a 10 mg/ml solution of α-cyano-4-hydroxycinnamic acid in ACN/50 mM citrate buffer, 70/30 v/v. Standard mixtures furnished by Applied Biosystems was used for calibrations of both MALDI mass spectrometers.

NanoLC-MS/MS experiments were performed on a 4000QTrap mass spectrometer (Applied Biosystems) coupled to a 1100 nanoHPLC system (Agilent Technologies). Peptide mixtures were loaded onto an Agilent reversed-phase pre-column cartridge (Zorbax 300 SB-C18, 5 x 75 µm, 3.5 µm) at 20 µl/min with solvent A (0.1 % formic acid, 2% ACN in water, loading time 5 min). Peptides were then separated on an Agilent reversed-phase column (Zorbax 300 SB-C18, 150 mm x 75 µm, 3.5 µm), at a flow rate of 300 nl/min using 0.1 % formic acid, 2% ACN in water as solvent A and 0.1% formic acid, 2% water in ACN as solvent B. The elution was accomplished by a 7-50% linear gradient of solvent B in 50 min. A nano-ionspray source was used at 2 kV with liquid coupling, with a declustering potential of 50 eV, using an uncoated silica tip (o.d. 150 µm, i.d. 20 µm, tip diameter 10 µm) from New Objectives (Ringoes, NJ). Spectra acquisition was based on a survey precursor ion scan for m/z 170. The Q_1 was scanned from 400 to 1400 in 2 s with resolution “low”, and the precursor ions were fragmented in q_2 using a linear gradient of collision energy from 30 to 80 eV. Finally, Q_3 was set to transmit only ions at m/z 170 with resolution “unit”. This scan mode was followed by an enhanced resolution experiment for the ions of interest and then by MS^2 and MS^3 experiments of the two most abundant species. MS^2 spectra were acquired using the best collision energy calculated on the basis of m/z values and charge state (*rolling collision energy*). MS^3 spectra were performed on the fragment ion at m/z 234, with a fixed LIT fill time of 350 ms and an activation time of 100 ms, scanning the mass range from 160 to 180 m/z.

III.3 Results and discussions

The analytical strategy described in this chapter fulfils the proteomics needs of sensitivity and specificity for the analysis of sub-stoichiometric post-translational modifications associated to protein oxidation [58], such as protein carbonylation. Compared to existing methodologies, essentially based on immunochemical

techniques [21-22,25], the strategy proposed here maintains sensitiveness but adds the capability to localise the modified residues; this information is often required to completely understand the biological significance of modification. 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 highly reproducible and can be used for RIGhT strategies [55-56].

III.3.1 Carbonylation of a model protein

The first part of this work concerns the optimization of an oxidation protocol able both to selectively oxidize lysine residues to α-aminoadipic semialdehyde residues, without undesired side reactions, and to introduce only a few numbers of modifications, in a controlled fashion, to reproduce *in vivo* conditions. This was necessary because of the commercial unavailability of carbonylated model proteins. The oxidation protocol was firstly optimized using a model peptide: the fragment 163-171 of the human interleukin-1β; this peptide has a molecular mass of 1004.40 Da (MH^+ = 1005.40 Da) and its primary structure is VQGEESNDK, with a single lysine residue at the C-terminus; for this reason this peptide was chosen as a model. Two different oxidation protocols were tested using NaClO [12] and $Fe(III)Cl_3$ [59]. Reaction conditions were optimized in terms of reagent concentrations, time and temperature. Using $Fe(III)Cl_3$ as oxidizing agent, the reaction produced a lot of undesired by-products not easy to characterize, while using NaClO the reaction was more selective and was easy to control, as demonstrated by the MALDI spectrum reported in figure III.3, where can be noted the presence of a most intense signal at m/z 1004.34, showing a ΔM = - 1 Da from the theoretical mass of the peptide, assigned to the oxidized form of the peptide. The ΔM of - 1 Da corresponds to the mass difference existing between lysine and α-aminoadipic semialdehyde after protein carbonylation event.

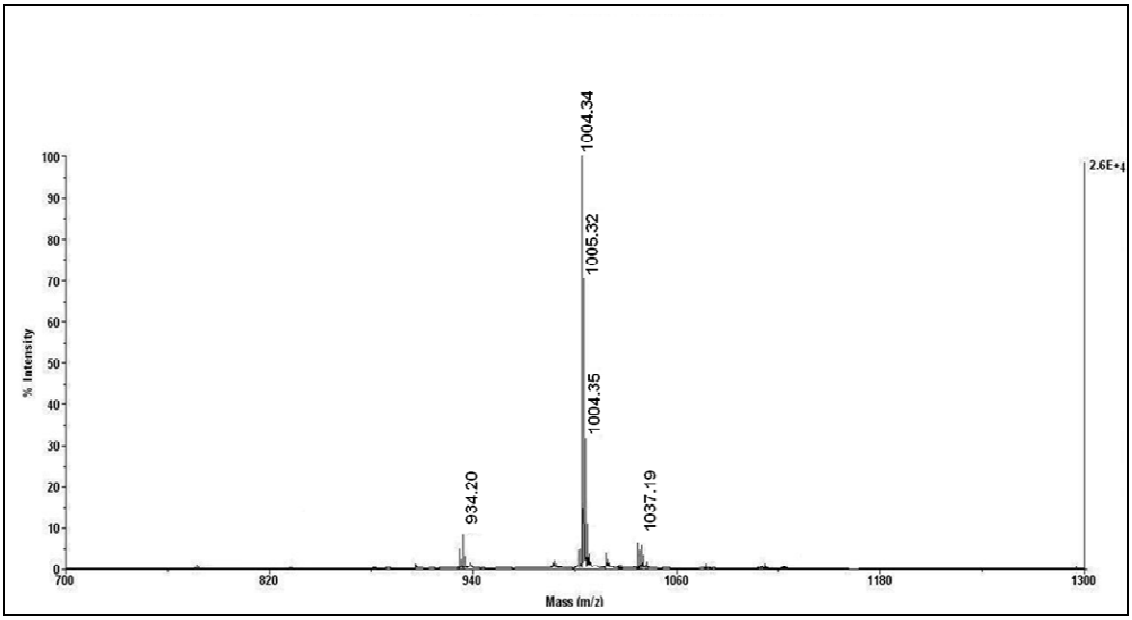


Figure III.3. *Particular of the MALDI spectrum of the oxidized interleukin-1β fragment; in the spectrum it can be noted the presence of a most intense signal at m/z 1004.34 that can be assigned to the oxidized form of the model peptide showing a ΔM = - 1 Da, corresponding to the mass difference existing between lysine and α-aminoadipic semialdehyde after protein carbonylation event.*

In fact, protein carbonyls generated by oxidation of lysine, arginine, or proline residues and frequently used as markers for oxidative protein modification, determine a difference in the molecular mass of amino acid side chains [60].While lysine residue side chain oxidation leads to the formation of α-aminoadipic semialdehyde, oxidation of proline and arginine residues side chain leads to the formation of glutamic semialdehyde. The hydroxyl group of threonine side chain can also be oxidized to form 2-amino-3-oxo-butanoic acid.
Reaction conditions for oxidation with NaClO were optimized as reported in material and methods section and the strategy was used for the oxidation of a model protein; the chosen proteins were bovine ribonuclease A and lysozyme, because of the fact that both proteins present a lysine residue at the N-terminus that should be easily accessible to oxidation mixture.
RNase A (0.5 mg) was oxidized by using NaClO under controlled conditions (see materials and methods) and, after reaction, rapidly purified by size exclusion chromatography to remove oxidising solution. Aliquots of the sample containing RNase A and carbonylated RNase A, were reduced with DTT, alkylated with IAM and digested with trypsin. The obtained peptide mixture was then analysed by MALDI mass spectrometry, in order to identify the modification(s) induced by NaClO treatment. MS analyses led to the verification of about 95% of protein sequence and let me to locate the carbonylated residue in the peptide 1-10. In fact, MALDI-MS spectra showed the occurrence of a signal at m/z 1149.63 (figure III.4) exhibiting a molecular mass not corresponding to a theoretical one within the RNase A sequence. This signal was assigned to the peptide 1-10 showing a ΔM = - 1 Da that, as described above, correspond to the conversion of a lysine residue to α-aminoadipic semialdehyde residue. The presence in the MALDI spectrum of the signal at m/z 1150.6 attributed to the theoretical peptide 1-10 of RNase A, indicates that the oxidation reaction was not quantitative; this represents a good result because it reproduces the sub-stoichiometric nature of this PTM. No evident signal occurring at m/z 1148.6, corresponding to the peptide 1-10 doubly modified by carbonylation at the second lysine residue, was detected.

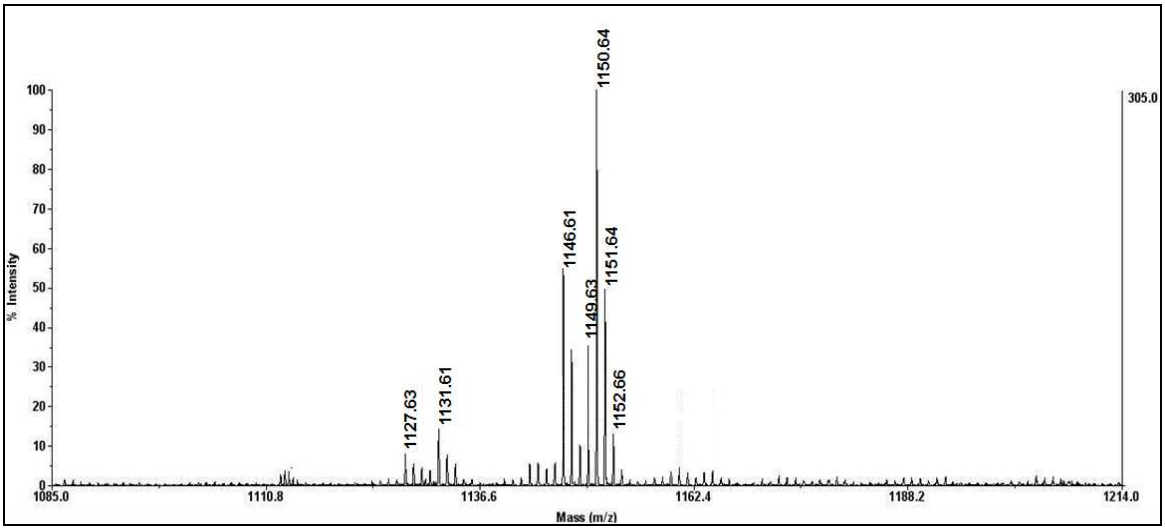


Figure III.4. *Particular of the MALDI spectrum of the oxidized RNase A; the spectrum shows a signal at m/z 1150.6 that can be attributed at the peptide 1-10 of the RNase A, while the signal at m/z 1149.6 can be assigned to the oxidized form of the same peptide (ΔM = - 1 Da). No evident signal occurring at m/z 1148.6, corresponding to the peptide 1-10 doubly modified by carbonylation at the second residue of lysine, can be detected.*

The same procedure was used for the oxidation of lysozyme. The tryptic mixture of oxidised lysozyme was then submitted to MALDI-MS analyses. The interpretation of MALDI-MS spectra revealed the occurrence of a signal whose molecular mass could not be explained on the basis of the theoretical masses. It should be noted that, as RNase A, lysozyme showed the presence of a lysine residue at the N-terminal position. In this case the expected signal at m/z 605.4 corresponding to carbonylated peptide 1-5 was not detected. However, a signal at m/z 587.4 (18 Da lower than carbonylated specie), was detected (figure III.5 A).

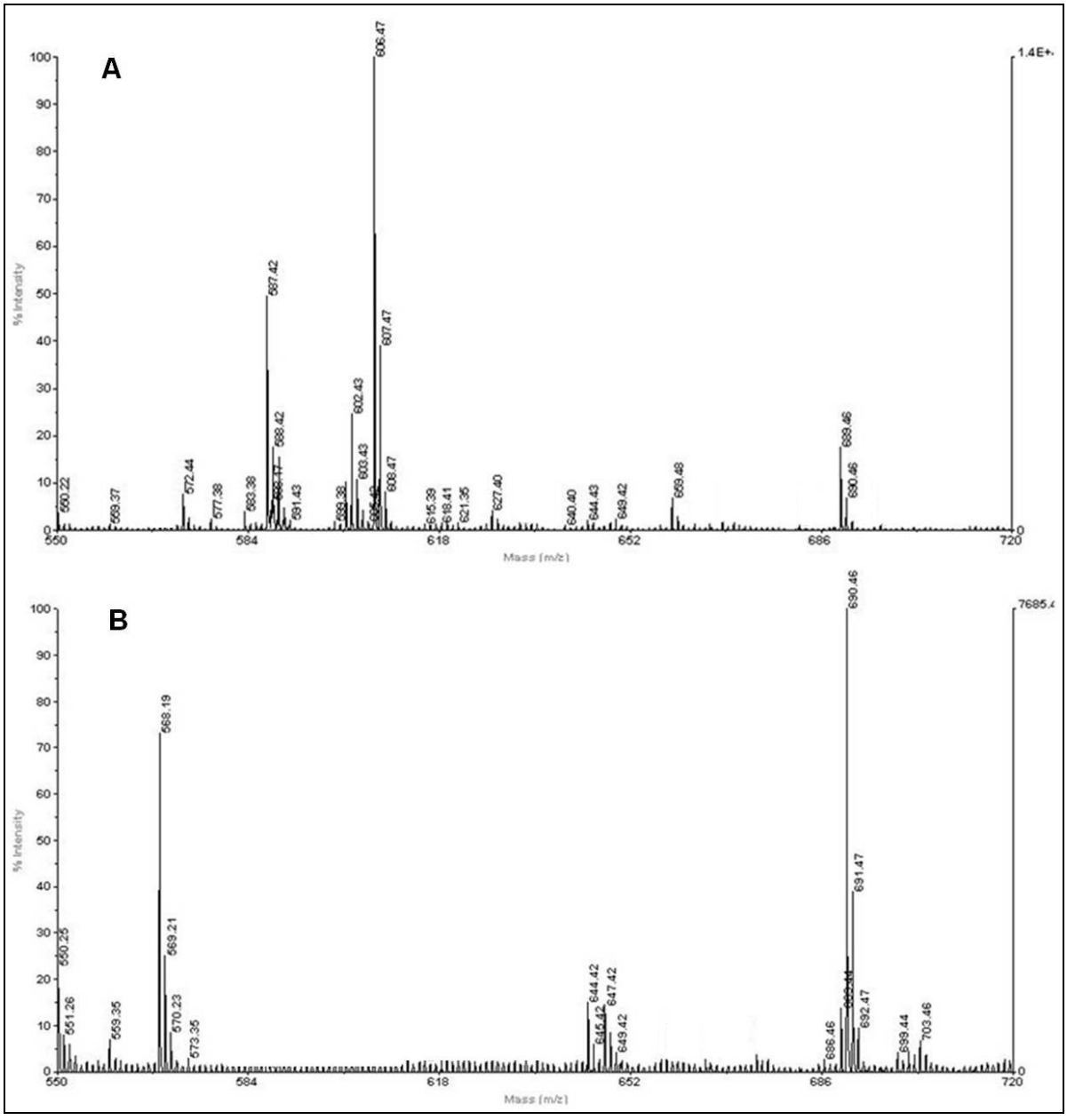


Figure III.5. *Particular of MALDI spectra of carbonylated lysozyme (panel A) and acetylated carbonylated lysozyme (panel B). In panel A spectrum is evident the presence of a signal at m/z 606.47, attributed to peptide 1-5 of lysozyme, and of a signal at m/z 587.42 tentatively assigned to a cyclic intermediate formed after carbonylation of protein. In panel B spectrum is evident the presence of two related signals at m/z 647.34 and 690.38. These signals were attributed to the peptide 1-5 carbonylated and acetylated (ΔM = + 42 Da) and the same peptide doubly acetylated, respectively. The absence of the signal at m/z 587.32 could be appreciated, thus confirming the formation of the cyclic intermediate at the N-terminal residue.*

This peak was absent in the spectrum of untreated lysozyme, and was tentatively assigned to a cyclic intermediate originated from the reaction of N-terminal amino group with the aldehyde of the oxidized lysine residue in position 1.
To support this hypothesis, the strategy was modified introducing an acetylation step to block amino groups (N-terminus and Lys ε-amino groups) [61].Following the reaction, the acetylated N-terminal amino group can not react with aldehyde moiety generated by oxidation. The acetylation reaction was carried out by using acetic acid N-hydroxysuccinimide esther (1000-fold molar excess on proteins). The mixtures were then analysed by MALDI mass spectrometry and the spectra revealed the disappearance of signals attributed to the cyclic intermediate formation. In fact the MALDI spectra showed the presence of two related signals at m/z 647.4 and 690.4, that were attributed to the peptide 1-5 carbonylated and acetylated (ΔM = + 42 Da) and to the same peptide doubly acetylated (ΔM = + 84 Da), respectively. The absence of the signal at m/z 587.4 (figure III.5 B) could be appreciated, thus confirming that acetylation of N-terminus amino group blocks the formation of the cyclic intermediate when α-aminoadipic semialdehyde is at the N-terminal residue.

III.3.2 Labelling with DNSH

The formation of the cyclic intermediate was an undesired side reaction since it blocks the reactivity of carbonylated lysine, which was not suitable for reaction with hydrazide moiety of DNS derivate. Moreover the acetylation of carbonylated lysozyme caused the introduction of a further step that complicated the analytical strategy and was not useful for my proposal. For these reasons RNase A was chosen as oxidised model protein. According to “gel-free procedures”, the analysis was carried out at level of peptides following tryptic digest of the whole protein mixture. Labelling of carbonylated residues of RNase A was carried out using dansyl-hydrazide (figure III.6).

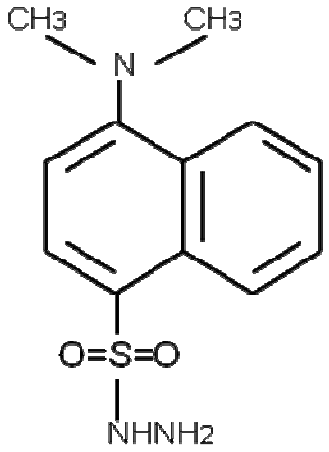


Figure III.6. *Dansyl-hydrazide molecule.*

The yield of the reaction was monitored by MALDI-MS analysis. Mass spectral analysis showed the disappearance of the signal at m/z 1149.6 (figure III.7 C), indicating the successful labelling of this specie. A new signal at m/z 1396.7 was detected showing a mass increment of + 247 Da (figure III.7 B). This signal was

attributed to the peptide 1-10 modified by α-aminoadipic semialdehyde hydrazone formation following dansyl-hydrazide reaction (figure III.7 A) on the RNase A.

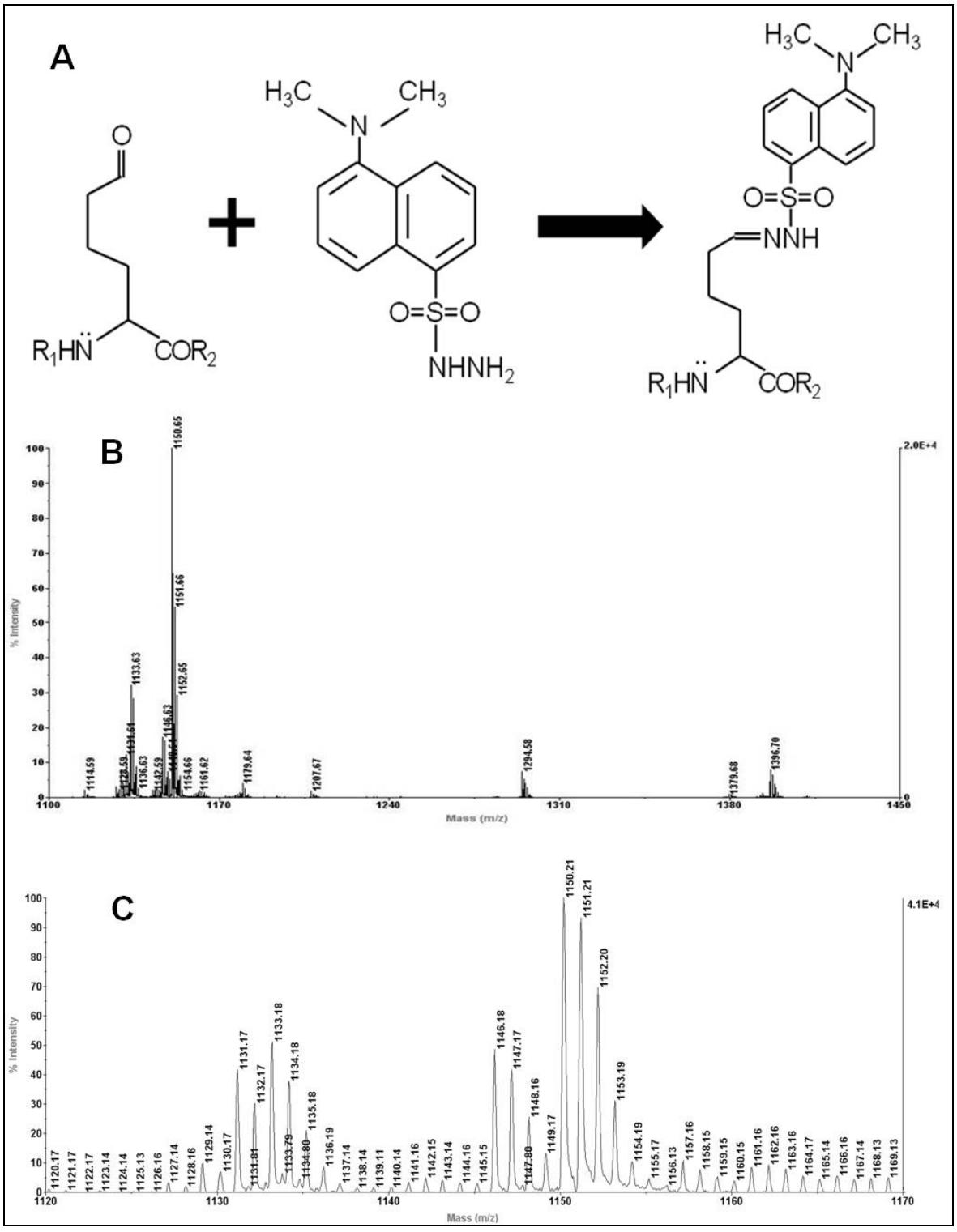


Figure III.7. *Panel A shows the formation of the α-aminoadipic semialdehyde hydrazone following dansyl-hydrazide reaction. Panel B shows the MALDI spectrum of the carbonylated RNase A peptide mixture after dansyl-hydrazide reaction. Mass spectral analysis showed the loss of the signal at m/z 1149.60, indicating the labelling of this specie. A new signal at m/z 1396.70 was detected showing a mass increment of + 247 Da. This signal was attributed to the peptide 1-10 modified by hydrazone formation (ΔM = + 247 Da). Panel C shows a particular of MALDI spectrum.*

Moreover the absence of further species with a ΔM = + 247 Da, indicated that the reagent was specific for carbonyl groups. MALDI-MS spectrum also showed the presence of a signal at m/z 1131.6 assigned to the cyclic intermediate originated from the reaction of N-terminal amino group with the aldehyde of the oxidized lysine residue in position 1; the formation of this intermediate caused the disappearance of carbonyl groups of lysine in position 1, thus this residue was not yet reactive towards hydrazide moiety.

III.3.3 LC-MS/MS analysis of dansylated peptides

Discrimination between modified and unmodified peptides was achieved by taking advantage of the instrumental features of a QqLIT mass spectrometer (section I.2.1). Analysis of peptide mixtures were carried out by using LC-MS/MS method; ions of interest were discriminated by a combination of two selection criteria, based on the combination of a precursor ion scan analysis as survey scan and an MS^3 scan, in a bidimensional tandem mass spectrometry selection [57]. By using precursor ion scan mode it was possible to selectively detect the only precursor ions producing the fragment at 170 m/z, generated by typical fragmentation of dansyl and its derivatives [56]. The selected precursor ions were then analyzed by combined MS^2 and MS^3 scan mode, to specifically select the dansylated species by exploiting the precursor ion scan selection for the fragment at 170 m/z and taking advantage of the dansyl-specific MS^3 transition m/z 234→170 (figure III.2). LC-MS/MS chromatograms obtained from the analysis of dansylated oxidised-RNase A peptide mixture are shown in figure III.8 A-B-C.

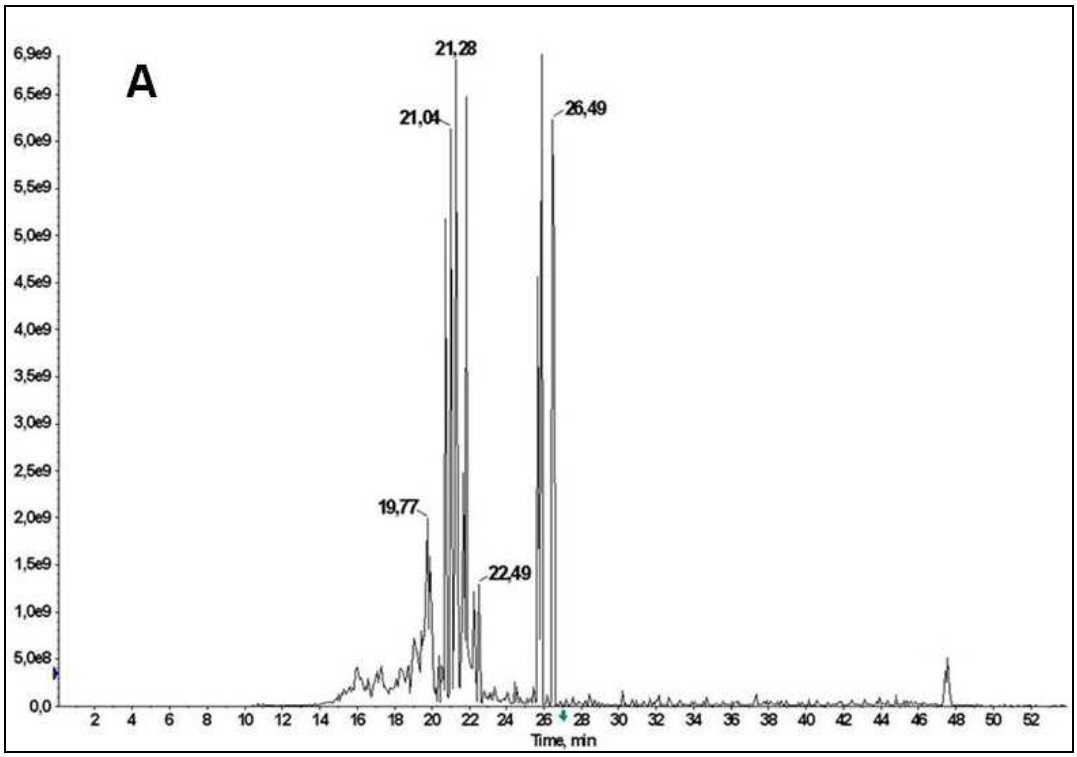


Figure III.8 A. *This panel shows the TIC of the Full Scan MS/MS analysis, in which are present a large number of signals most of which not related to dansyl modified peptides.*

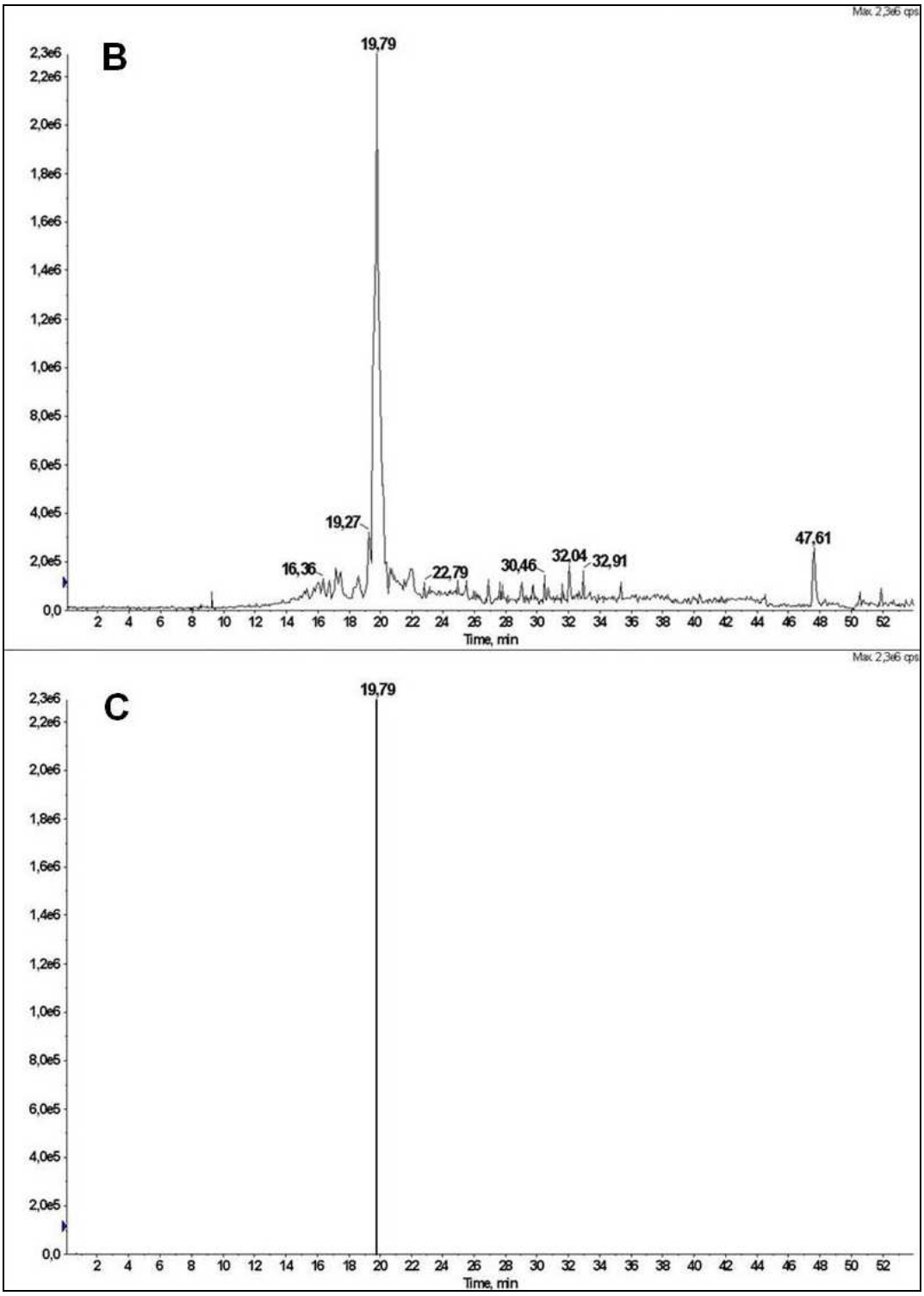


Figure III.8 B - C. *The combination of PIS and MS^3 (panels B and C) analysis let me to obtain a significant increase in signal/noise ratio and a significant increase in sensitivity of mass spectrometric analysis. The unique signal at 19.8 min in PIS TIC was correlated to 1-10 peptide of RNase A carbonylated and labelled by dansyl-hydrazide.*

By comparing the TIC profiles from PIS and MS^3 (panels B and C) with the TIC chromatogram of LC-MS/MS full scan analysis (panel A), the precursor ion scan mode still showed the occurrence of a large number of signals, most of which not related to dansyl modified peptides. In fact, the corresponding MS^2 spectra did not

show the occurrence of fragment ion at m/z 234. The further selection based on the MS^3 scan removed a large number of false positives leading to a simple ion chromatogram essentially dominated by a unique signal, indicating the presence of a single dansyl-modified peptides.
The combination of PIS and MS^3 analysis let me to obtain a significant increase in signal/noise ratio and a significant increase in sensitivity of mass spectrometric analysis, due to the reduction of duty cycle of the mass spectrometer. The unique peak with a retention time of 19.8 min in PIS chromatogram was correlated by its MS and MS/MS spectra to peptide 1-10 of carbonylated RNase A, labeled by dansyl-hydrazide. Manual interpretation of fragmentation spectrum obtained from this species, allowed to identify the modified peptide and to locate the modified residue. Figure III.9 shows the MS^2 spectrum of triply charged ion at m/z 466.2 attributed to peptide 1-10 of RNase A carrying a dansyl-hydrazone moiety. Modified lysine residue was easily detected by interpretation of fragmentation spectrum, thus confirming previous hypothesis of carbonylation at level of lysine residue in position 7 of this peptide. This MS^2 spectrum showed the presence of fragment ions belonging to both the y and the b series. However, some features made difficult MS/MS spectrum interpretation. In particular the modified peptide was not very stable during collision induced dissociation; moreover the peptide bond between alanine in position 6 and modified α-aminoadipic semialdehyde in position 7 was not completely fragmented under condition used in the experiment. Thus the occurrence of a signal corresponding to y_4 ion could be detected even if at very low intensity. This fragment ion corresponded to the peptide sequence KFER still containing lysine residue modified by dansyl moiety. Moreover the neutral loss of DNS moiety (m/z 234 Da) from modified α-aminoadipic semialdehyde residue could be appreciated. In fact it could be observed y ions exhibiting a ΔM of 233 Da, corresponding to y fragments that have lost DNS moiety but that still retained the hydrazone moiety (figure III.9).

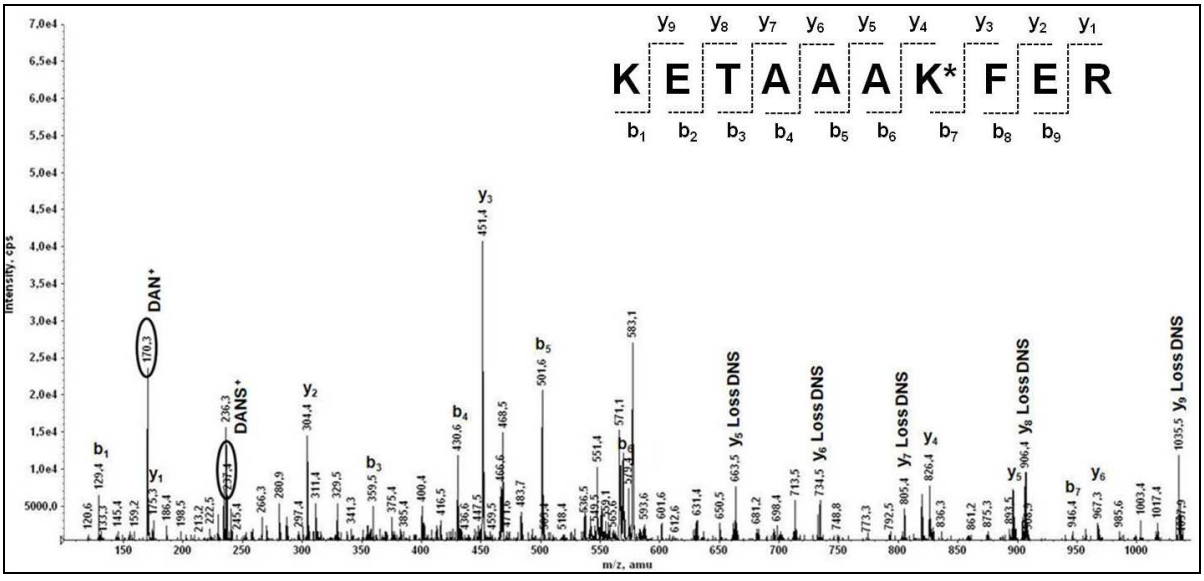


Figure III.9. *MS^2 spectrum of triply charged ion at m/z 466.2 of the peptide 1-10 carrying a dansyl-hydrazone moiety.*

III.3.4 Analysis of a complex protein mixture

The proposed strategy was finally applied to the analysis of a more complex sample, in order to demonstrate the ability of this analytical method to specifically locate carbonylated residues on a proteomics scale. For this purpose 100 μg of the carbonylated RNase A were spiked with 10 mg of an entire protein extract of *E. coli* cells. This complex mixture was analysed using the procedure described above. Disulfide bridges were reduced by DTT and free cysteines were alkylated by IAM; then protein mixture was digested with trypsin and the resulting peptide mixture was submitted to labelling with dansyl-hydrazide as described. An aliquot of the resulting dansylated peptide mixture was submitted to bidimensional LC-MS/MS analysis in PIS and MS^3 modes as described above. Results of LC-MS/MS analysis are summarized in figure III.10, showing the chromatographic profiles of TIC (panel A) and PIS (panel B) experiments. In the TIC profile a large number of signals are present, that could be due both to false positives or real dansylated peptides.
However the selectivity of the bidimensional analysis allowed to eliminate all false positives since by MS^3 analysis a unique signal was detected, attributable to carbonylated and dansylated peptide 1-10 of RNase A. Manual interpretation of MS^2 spectrum related to this signal allowed me to reconstruct the sequence of the above mentioned peptide and to locate the modification site on the lysine in position 7.To the best of my knowledge this strategy represents the first method leading to the direct identification of carbonylation sites in proteins, thus indicating the feasibility of this strategy to investigate protein carbonylation in proteomics approaches. However by using dansyl-hydrazide it is not possible to perform quantitative analysis of carbonylation sites. For this reason it was necessary to design a new strategy that makes use of an iTRAQ derivative for specific labelling and quantitative analysis of carbonyl groups of proteins.

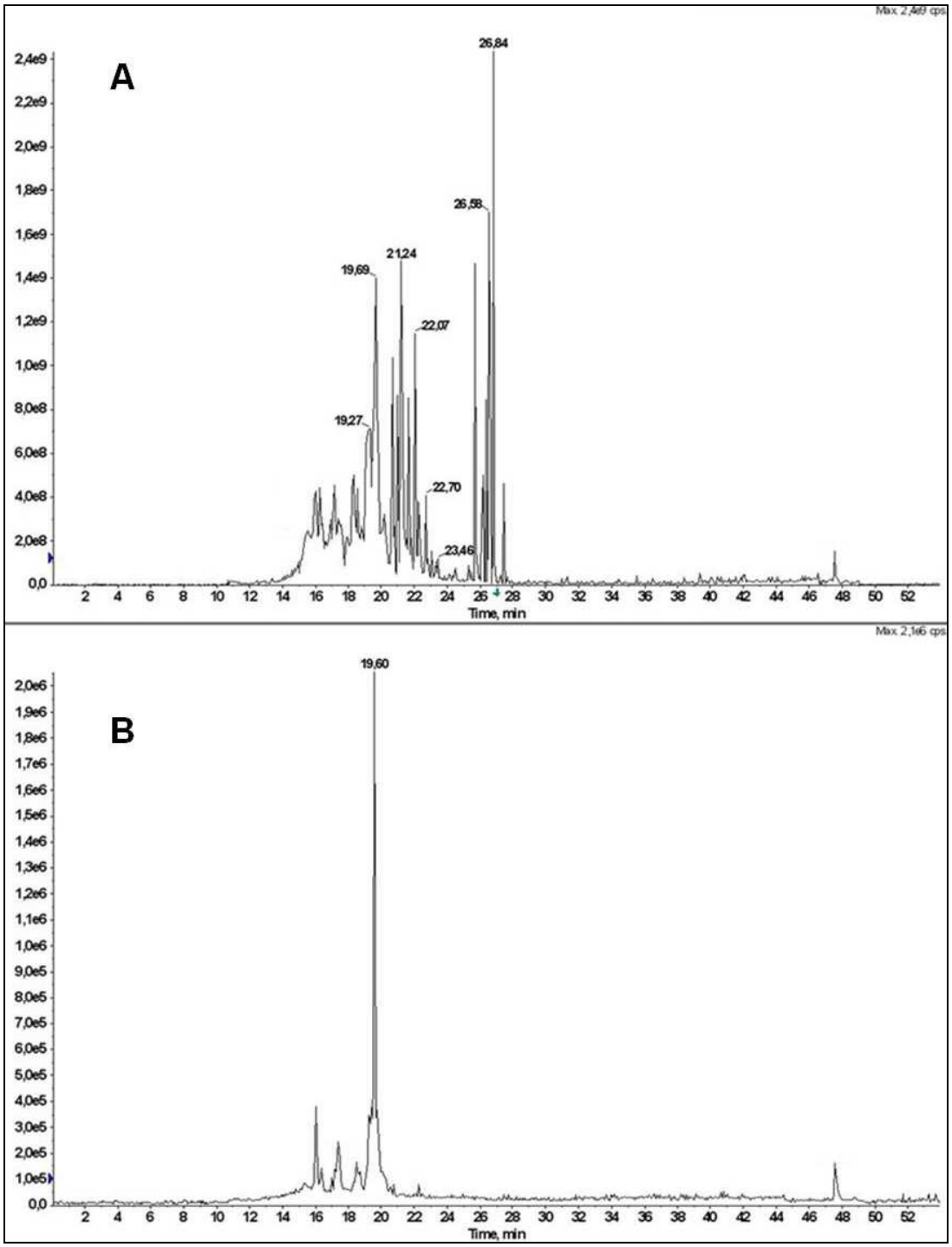


Figure III.10. *Chromatographic profiles of PIS (panel A) and MS^3 (panel B) experiments obtained by bidimensional LC-MS/MS analysis of 100 μg of the carbonylated RNase A was spiked to 10 mg of entire cellular extract of E. coli. The precursor ion scan profile still shows a large number of signals that could be due both to false positives or real dansylated peptides. However the selectivity of the MS^3 analysis allowed us to eliminate all false positives since in the MS^3 profile a unique signal referable to carbonylated and dansylated peptide 1-10 of RNase A was detected.*

III.3.5 iTRAQ-hydrazide labelling of standard mixtures

iTRAQ molecules are specific toward amino groups, so in order to use iTRAQ for the specific labelling of carbonylated lysine residues, it was necessary to modify iTRAQ

molecule. For this purpose it was designed a strategy to obtain a new iTRAQ derivative, namely iTRAQ-hydrazide (iTRAQH), by reacting iTRAQ and hydrazine molecules (figure III.11).

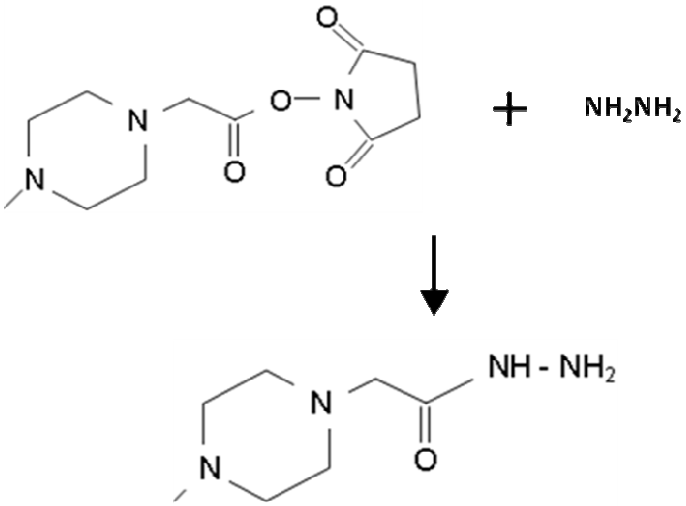


Figure III.11. *Reaction scheme of the synthesis of iTRAQ-hydrazide molecule.*

The synthesis of the iTRAQ-hydrazide molecule was achieved by reacting iTRAQ and hydrazine following the procedure described in materials and methods section. Reaction product was purified by RP-HPLC and analysed by MALDI mass spectrometry in order to verify the success of reaction. As shown in the MALDI spectrum reported in figure III.12, a signal at m/z 177.2 is present, which corresponds to the expected m/z of the iTRAQ-hydrazide product (molecular mass = 176.2 Da).
As for DNSH labelling, the strategy was optimised using oxidised RNase A model protein. According to “gel-free procedures”, the labelling was carried out at level of peptides following tryptic digestion of the whole protein mixture. An aliquot of carbonylated RNase A was reduced by DTT and free thiols were alkylated by IAM prior to tryptic digestion.
A solution of iTRAQ-hydrazide freshly synthesized was added to the oxidised peptide mixture and the reaction was carried out as reported in the materials and methods section, following a slightly different strategy optimized for DNSH labelling.

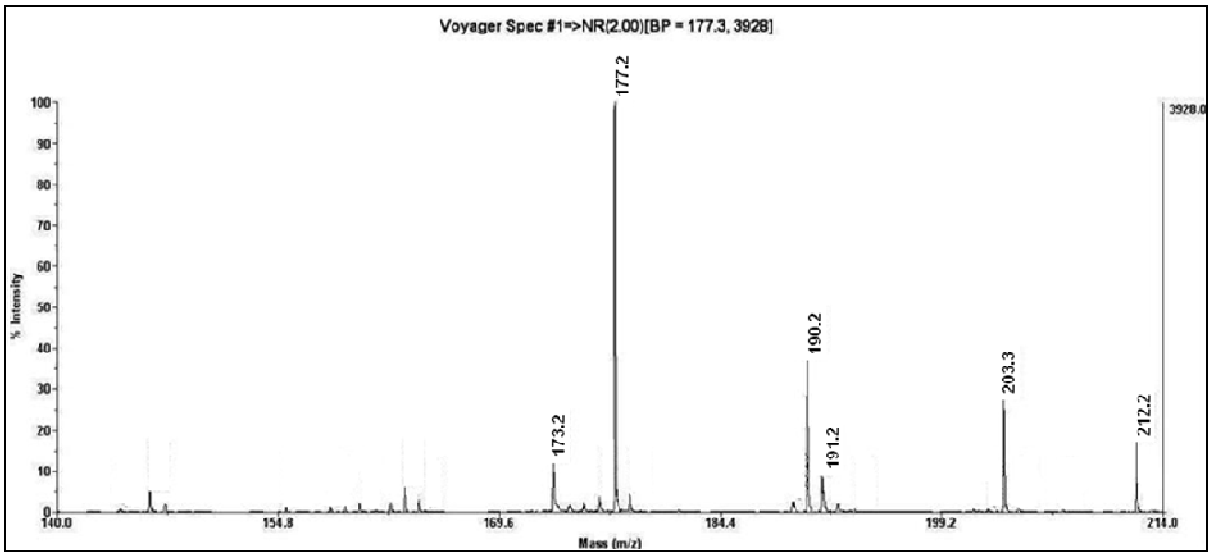


Figure III.12. *Particular of the MALDI spectrum obtained from the analysis of iTRAQ-hydrazide newly synthesised. The spectrum shows a signal at m/z 177.2 that corresponds to the expected m/z value of the iTRAQ-hydrazide molecule.*

After reaction the modified peptide mixture was analysed by MALDI mass spectrometry in order to verify the success of the labelling. Figure III.13 shows the MALDI spectrum obtained from the analysis of modified peptide mixture; in the spectrum is still present the signal at m/z 1149.6 attributable to the carbonylated peptide 1-10 of RNase A, but is also present a signal at m/z 1307.7 (ΔM = + 158 Da) assigned to the carbonylated peptide 1-10 of RNase A modified with iTRAQ-hydrazide, resulting in the formation of an hydrazone moiety of the iTRAQ derivative.

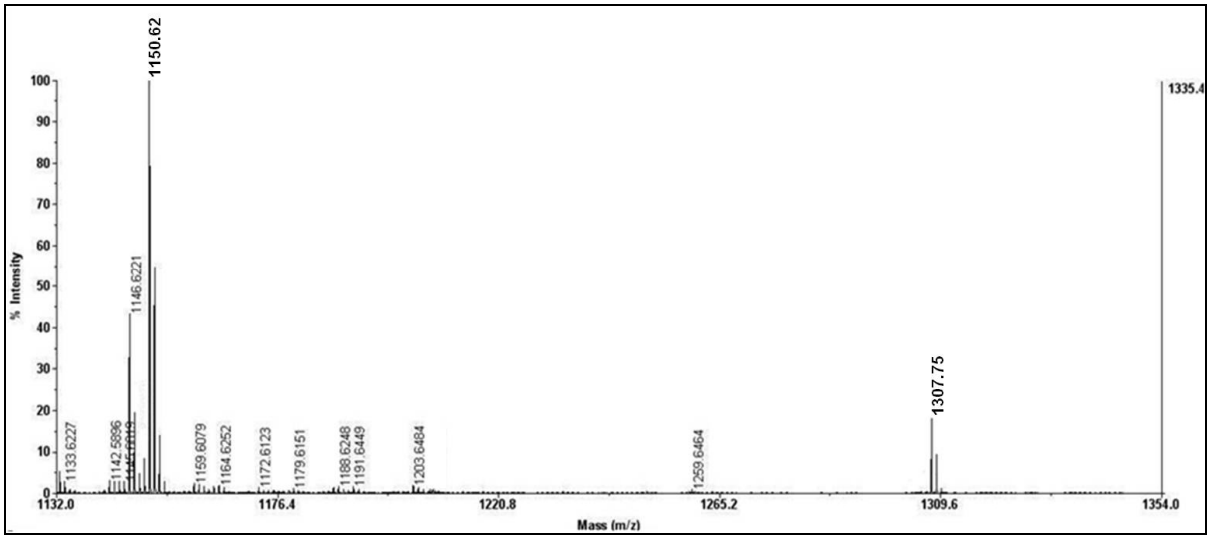


Figure III.13. *Particular of the MALDI spectrum obtained from the analysis of iTRAQ-hydrazide labelled peptide mixture of oxidized RNase A. In this spectrum is present a signal at m/z 1307.7 (ΔM = + 158 Da) attributed to the carbonylated peptide 1-10 of RNase A modified with iTRAQ-hydrazide with the formation of a hydrazone moiety of the iTRAQ derivative.*

The use of iTRAQ-hydrazide to label oxidised lysine residues makes possible to quantitate differentially oxidised proteins; for this purpose two different iTRAQ molecules were chosen to synthesize the hydrazide derivatives. In particular iTRAQ

114 and iTRAQ 117 were used and it was possible to obtain iTRAQ-hydrazide 114 (iTRAQH 114) and iTRAQ-hydrazide 117 (iTRAQH 117).
Aliquots of peptide mixtures of oxidised RNase A, labelled with the two different isobaric variants of iTRAQ-hydrazide, were mixed in order to obtain two different mixtures in which molar ratios between RNase A labelled with iTRAQH 114 and RNase A labelled with iTRAQH 117 were 1:1 and 1:4, respectively. However because of the isobaric nature of iTRAQ molecules (section II.1.2), differentially labelled peptides are indistinguishable by a simple MS analysis, whereas the iTRAQ reporter ions are generated after fragmentation events, thus allowing the quantitative analysis by using MS/MS analysis. The two peptide mixtures were analyzed by using a MALDI-TOFTOF mass spectrometer in order to verify the feasibility of the quantitative approach. Among the MS/MS spectra obtained from the analysis of the peptide mixture, it was possible to find iTRAQ reporter signals only in one spectrum as expected; this spectrum was originated by the precursor ion at m/z 1307.7 attributed to the oxidised peptide 1-10 of RNase A labelled by iTRAQ-hydrazide (figure III.14).
By manual interpretation of MS/MS spectrum it was possible to reconstruct the sequence of the peptide 1-10 of RNase A both with y ions and b ions retaining the modified residue, thus indicating a stable modification of the residue of interest; moreover it was possible to identify the modification site on the lysine residue in position 7 as for the case of DNSH modified protein. In the low mass region of the spectrum it was also possible to find the signals of iTRAQ reporter ions that could be used for quantitative analysis.

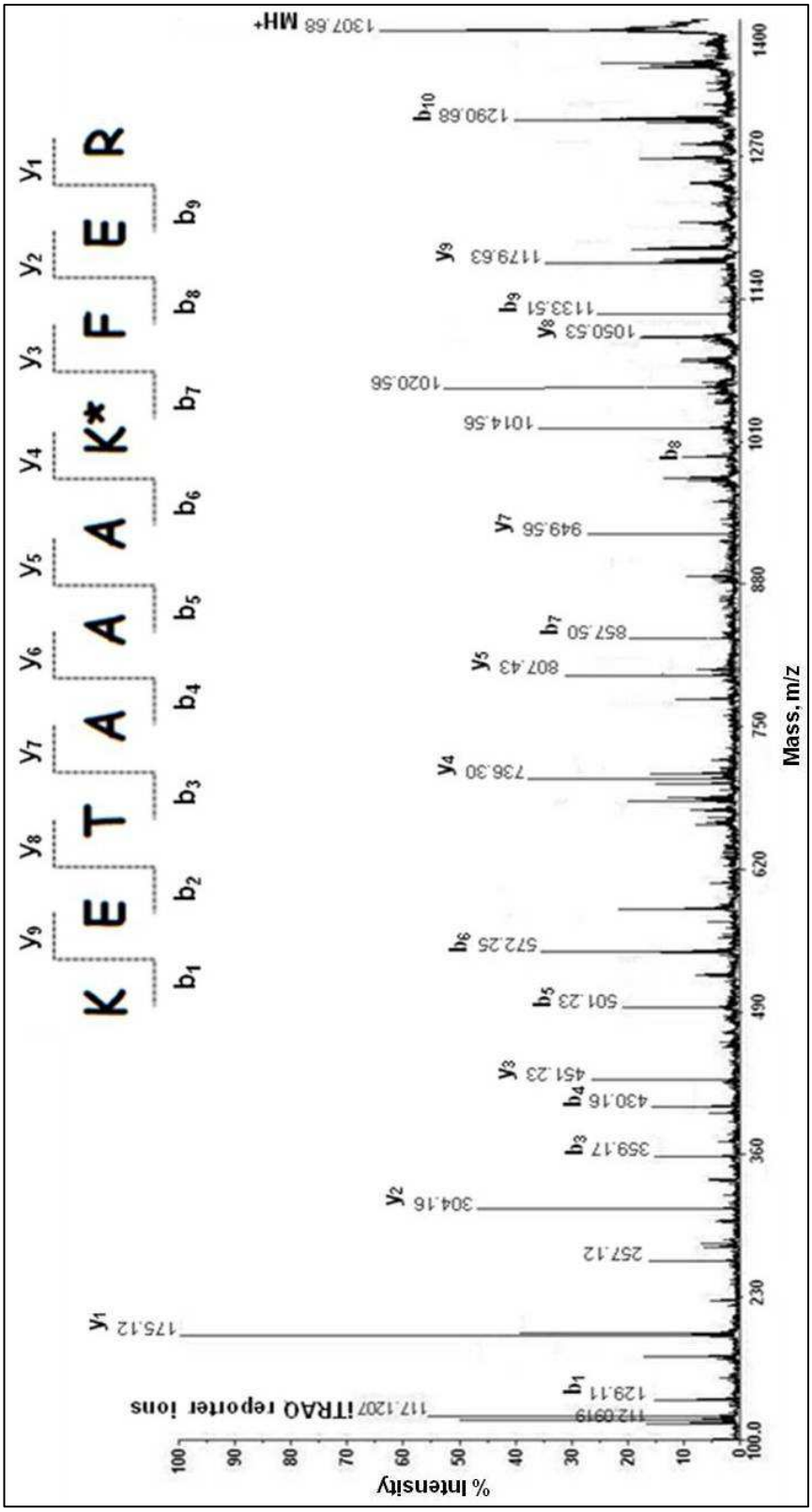


Figure III.14. *MS/MS spectrum of the precursor ion 1307.7 attributed to iTRAQ modified oxidized peptide 1-10 of RNase A. articular of the MALDI spectrum obtained from the analysis of iTRAQ-hydrazide labelled peptide mixture of oxidized RNase A. In the spectrum are present all the expected fragments of y and b series and by manual interpretation it was possible to reconstruct the entire sequence of the peptide 1-10 of RNase A retaining the modified residue; moreover it was possible to specifically identify the modification site on the lysine residue in position 7. The low mass region of the spectrum shows the signals of iTRAQ reporter ions that could be used for quantitative analysis.*

The two mixtures (1:1 and 1:4) were analysed separately and it was possible to determine the molar ratios of the two differentially labelled peptide mixtures by measuring the area above the signals of iTRAQ reporter ions. Figure III.15 shows a magnified section of an MS/MS spectrum for both mixtures, in which are shown the signals of iTRAQ. reporter ions By replicating these experiments it was possible to calculate an experimental ratio of 1.07 ± 0.03 for the 1:1 mixture and 3.97 ± 0.17 for the 1:4 mixture. These preliminary data demonstrate the feasibility of this method for quantitative analysis of protein carbonylation.

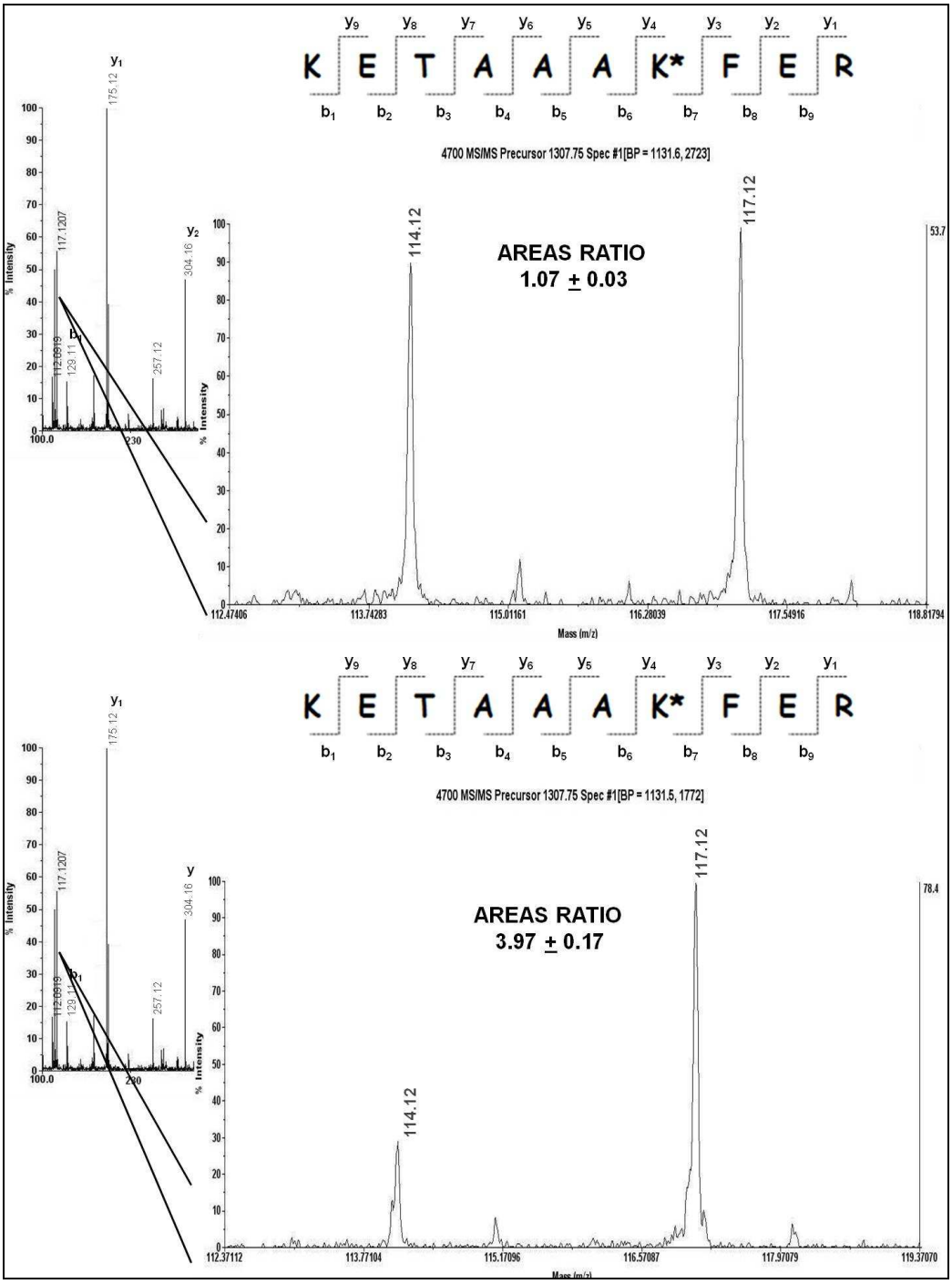


Figure III.15. *Magnified sections of MS/MS spectra of modified peptide 1-10 of RNase A for 1:1 and 1:4 mixtures. The signals of reporter ions of iTRAQ can be used for quantitative analysis.*

III.4 Conclusions

In this chapter the use of dansyl-hydrazide in combination with advanced mass spectrometric experiments, allowed the selective identification of carbonylation sites of RNase A also in a complex mixture as the entire protein extract of *E. coli*. The generation of carbonyl groups in RNase A was obtained by the oxidation with NaClO, with a slight modification of the protocol proposed by Headlam and co-workers [12]; it was possible to specifically convert just one lysine residue (namely Lys 7) of RNase A to α-aminoadipic semialdehyde, thus miming the *in vivo* conditions. Dansyl-hydrazide was then used to specifically label α-aminoadipic semialdehyde residues and the peculiar fragmentation pathway of dansyl derivatives was then exploited for mass spectrometric analysis with a hybrid instrument, such as QqLIT, in *Precursor Ion Scan* analysis, combined with MS^3 analysis, in a bidimensional LC-MS/MS method. The combination of PIS and MS^3 let me to obtain a significant increase in signal/noise ratio and a significant increase in sensitivity of mass spectrometric analysis, due to the reduction of duty cycle of the mass spectrometer. By this way it was possible to specifically identify the only modified peptide of RNase A and, at the same time, to locate the modification sites (the oxidised residue). This work represents the first method leading to the direct identification of carbonylation sites in proteins; moreover the capability of bidimensional mass spectrometric analysis to selectively identify the modified peptide in a complex peptide mixture indicates the feasibility of this strategy to investigate protein carbonylation on a proteomics scale.
However by using dansyl-hydrazide it was not possible to specifically quantify the extent of protein carbonylation; for this reason a new iTRAQ derivative was synthesized exploiting the specificity of hydrazide derivatives toward carbonyl groups in combination with the unique features of iTRAQ reagents for quantitative analysis; this new iTRAQ-hydrazide reagent was used to label a tryptic peptide mixture of oxidised RNase A. This strategy has to be optimized for complex mixtures analysis, but initial results led to both selective qualitative and quantitative analysis of carbonylated residues.

III.5 References

1. Shacter, E. *Methods Enzymol*. **2000**, *319*, 428-436.
2. Stadtman, E.R. *Ann NY Acad Sci*. **2001**, *928*, 22-38.
3. Sorolla, M.A.; Reverter-Branchat, G.; Tamarit, J.; Ferrer, I.; Ros, J.; Cabiscol, E. *Free Radic Biol Med*. **2008**, *45*, 667-678.
4. Yoo, B.S.; Regnier, F.E. *Electrophoresis*. **2004**, *25*, 1334-1341.
5. Magi, B.; Ettorre, A.; Liberatori, S.; Bini, L.; Andreassi, M.; Frosali, S.; Neri, P.; Pallini, V.; Di Stefano, A. *Cell Death Differ*. **2004**, *11*, 842-852.
6. Johansson, E.; Olsson, O.; Nyström, T. *J Biol Chem*. **2004**, *279*, 22204-22208.
7. England, K.; O'Driscoll, C.; Cotter, T.G. *Cell Death Differ*. **2004**, *11*, 252-260.
8. England, K.; Cotter, T.G. *Biochem Biophys Res Commun*. **2004**, *320*, 123-130.
9. Soreghan, B.A.; Yang, F.; Thomas, S.N.; Hsu, J.; Yang, A.J. *Pharm Res*. **2003**, *20*, 1713-1720.
10. Levine, R.L.; Wehr, N.; Williams, J.A.; Stadtman, E.R.; Shacter, E. *Methods Mol Biol*. **2000**, *99*, 15-24.

11. Ghezzi, P.; Bonetto, V. *Proteomics*. **2003**, *3*, 1145-1153.
12. Headlam, H.A.; Davies, M.J. *Free Radic Biol Med*. **2004**, *36*, 1175-1184.
13. Madian, A.G.; Regnier, F.E. *J Proteome Res*. **2010**, *9*, 1330-1343.
14. Mirzaei, H.; Regnier, F. *Anal Chem*. **2005**, *77*, 2386–2392.
15. Mirzaei, H.; Regnier, F. *Anal Chem*. **2006**, *78*, 770–778.
16. Levine, R.L.; Garland, D.; Oliver, C.N.; Amici, A.; Climent, I.; Lenz, A.G.; Ahn, B.W.; Shaltiel, S.; Stadtman, E.R. *Methods Enzymol*. **1990**, *186*, 464-478.
17. Levine, R.L.; Williams, J.A.; Stadtman, E.R.; Shacter, E. *Methods Enzymol*. **1994**, *233*, 346-357.
18. Yan, L.J. *Curr Protoc Protein Sci*. **2009**, *Chapter 14*, Unit 14.4.
19. Keller, R.J.; Halmes, N.C.; Hinson, J.A.; Pumford, N.R. *Chem Res Toxicol*. **1993**, *6*, 430-433.
20. Shacter, E.; Williams, J.A.; Lim, M.; Levine, R.L. *Free Radic Biol Med*. **1994**, *17*, 429-437.
21. Sharma, R.; Nakamura, A.; Takahashi, R.; Nakamoto, H.; Goto, S. *Free Radic Biol Med*. **2006**, *40*, 1179-1184.
22. Blackburn, A.C.; Doe, W.F.; Buffinton, G.D. *Free Radic Biol Med*. **1999**, *27*, 262-270.
23. Yan, L.J.; Levine, R.L.; Sohal, R.S. *Proc Natl Acad Sci USA*. **1997**, *94*, 11168-11172.
24. Yan, L.J.; Sohal, R.S. *Proc Natl Acad Sci USA*. **1998**, *95*, 12896-12901.
25. Nakamura, A.; Goto, S. *J Biochem*. **1996**, *119*, 768-774.
26. Choi, J.; Forster, M.J.; McDonald, S.R.; Weintraub, S.T.; Carroll, C.A.; Gracy, R.W. *Free Radic Biol Med*. **2004**, *36*, 1155-1162.
27. Boyd-Kimball, D.; Castegna, A.; Sultana, R.; Poon, H.F.; Petroze, R.; Lynn, B.C.; Klein, J.B.; Butterfield, D.A. *Brain Res*. **2005**, *1044*, 206-215.
28. Mello, C.F.; Sultana, R.; Piroddi, M.; Cai, J.; Pierce, W.M.; Klein, J.B.; Butterfield, D.A. *Neuroscience*. **2007**, *147*, 674-679.
29. Reverter-Branchat, G.; Cabiscol, E.; Tamarit, J.; Ros, J. *J Biol Chem*. **2004**, *279*, 31983-31989.
30. Barreiro, E.; Gea, J.; Di Falco, M.; Kriazhev, L.; James, S.; Hussain, S.N. *Am J Respir Cell Mol Biol*. **2005**, *32*, 9-17.
31. Aguilaniu, H.; Gustafsson, L.; Rigoulet, M.; Nystrom, T. *Science*. **2003**, *299*, 1751-1753.
32. Desnues, B.; Cuny, C.; Gregori, G.; Dukan, S.; Aguilaniu, H.; Nystrom, T. *EMBO Rep*. **2003**, *4*, 400-404.
33. Hernebring, M.; Brolen, G.; Aguilaniu, H.; Semb, H.; Nystrom, T. *Proc Natl Acad Sci USA*. **2006**, *103*, 7700-7705.
34. Erjavec, N.; Larsson, L.; Grantham, J.; Nystrom, T. *Genes Dev*. **2007**, *21*, 2410-2421.
35. Frank, J.; Pompella, A.; Biesalski, H.K. *Free Radic Biol Med*. **2000**, *29*, 1096-1105.
36. Frank, J.; Pompella, A.; Biesalski, H.K. *Methods Mol Biol*. **2002**, *196*, 35-40.

37. Chaudhuri, A.R.; deWaal, E.M.; Pierce, A.; Van Remmen, H.; Ward, W.F.; Richardson, A. *Mech Ageing Dev*. **2006,** *127*, 849.
38. Jung, Y.; Chay, K.; Song, D.; Yang, S.; Lee, M.; Ahn, B.; *Arch Biochem Biophys*. **1997**, *345*, 311-317.
39. Ahn, B.; Rhee, S.G.; Stadtman, E.R. *Anal Biochem*. **1987**, *161*, 245-257.
40. Poon, H.F.; Abdullah, L.; Reed, J.; Doore, S.M.; Laird, C.; Mathura, V.; Mullan, M.; Crawford, F. *Biol Proced Online*. **2007**, *9*, 65-72.
41. Langhals, H.; Jona, W.; *Chem Eur J*. **1998**, *4*, 2110.
42. Lenz, A.G.; Costabel, U.; Shaltiel, S.; Levine, R.L. *Anal Biochem*. **1989**, *177*, 419-425.
43. Yan, L.J.; Sohal, R.S. *Anal Biochem*. **1998**, *265*, 176-182.
44. Gitlin, G.; Bayer, E.A.; Wilchek, M. *Biochem J*. **1987**, *242*, 923-926.
45. Temple, A.; Yen, T.Y.; Gronert, S. *J Am Soc Mass Spectrom*. **2006**, *17*, 1172-1180.
46. Mirzaei, H.; Regnier, F. *J Chromatogr A*. **2007**, *1141*, 22-31.
47. Meany, D.L.; Xie, H.; Thompson, L.V.; Arriaga, E.A.; Griffin, T.J. *Proteomics*. **2007**, *7*, 1150-1163.
48. Mirzaei, H.; Baena, B.; Barbas, C., Regnier, F. *Proteomics*. **2008**, *8*, 1516-1527.
49. Hensley, K. *Methods Mol Biol*. **2009**, *536*, 457-462.
50. Caperna, T.J.; Shannon, A.E.; Blomberg Le, A.; Garrett, W.M.; Ramsay, T.G. *Comp Biochem Physiol B Biochem Mol Biol*. **2010**, *156*, 189-196.
51. Soreghan, B.A.; Lu, B.W.; Thomas, S.N.; Duff, K.; Rakhmatulin, E.A.; Nikolskaya, T.; Chen, T.; Yang, A.J. *J Alzheimers Dis*. **2005**, *8*, 227-241.
52. McDonagh, B.; Ogueta, S.; Lasarte, G.; Padilla, C.A.; Barcena, J.A. *J Proteomics*. **2009**, *72*, 677-689.
53. Roe, M.R.; Xie, H.; Bandhakavi, S.; Griffin, T.J. *Anal Chem*. **2007**, *79*, 3747-3756.
54. Chavez, J.; Chung, W.G.; Miranda, C.L.; Singhal, M.; Stevens, J.F.; Maier, C.S. *Chem Res Toxicol*. **2010**, *23*, 37-47.
55. Amoresano, A.; Monti, G.; Cirulli, C.; Marino, G. *Rapid Commun Mass Spectrom*. **2006**, *20* (9), 1400-1404.
56. Amoresano, A.; Chiappetta, G.; Pucci, P.; D'Ischia, M.; Marino, G. *Anal Chem*. **2007**, *79* (5), 2109-2117.
57. Marino, G.; Buonocore, V. *Biochem J*. **1968**, *110*, 603-604.
58. Nakamura, A.; Kawakami, K.; Kametani, F.; Nakamoto, H.; Goto, S. *Ann NY Acad Sci*. **2010**, *1197*, 33-39.
59. Linert, W.; Jameson, G.N. *J Inorg Biochem*. **2000**, *79* (1-4), 319-326.
60. Requena, J.R.; Chao, C.C.; Levine, R.L.; Stadtman, E.R. *Proc Natl Acad Sci USA*. **2001**, *98*, 69-74.
61. Brancia, F.L.; Oliver, S.G.; Gaskell, SJ. *Rapid Commun Mass Spectrom*. **2000**, *14*, 2070-2073.

IIII. QUANTITATIVE IDENTIFICATION OF PROTEIN PHOSPHORYLATION SITES

IIII.1 Introduction

This chapter focuses on the development of a new proteomics approach aimed to study and quantify the extent of protein phosphorylation. In order to develop an analytical methodology that could be classified as RIGhT (Reporter Ion Generating Tag) (section I.3.1), iTRAQ reagents were chosen as the starting molecules for labelling of phosphorylation sites. At the best of my knowledge the iTRAQ chemistry has been limited so far to primary amines [1-5]. This work lies within a research line of our laboratory intended to expand the chemistry of this class of reagents with the aim to quantify PTMs carrying chemical moieties different from primary amines.

IIII.1.1 Proteomics analysis of phosphorylated proteins

Analysis of protein phosphorylation at a proteomics level still represents a challenging and highly relevant task both for the impact of enzyme-catalyzed phosphorylation and dephosphorylation on the function of many proteins [6] and for the sub-stoichiometric nature of this post-translational modification [7]; usually while some residues are always phosphorylated, others may only be transiently phosphorylated; moreover phosphoproteins abundance is very low and it can be quantified as 1-2% of the total protein amount in a cell. For these reasons it is necessary to develop analytical strategies able to isolate or enrich phosphorylated species before identification in order to increase the sensitivity and the selectivity of the analysis.
Protein phosphorylation has been traditionally studied and quantified by using radioactive labelling with ^{32}P followed by 2D-GE [8-9] or Edman degradation chemistry to sequence phosphopeptides in order to locate phosphorylation site(s) [6]. Both techniques have the advantage to allow the contemporary detection of N-, S- and O-phosphorylation. Quantitative analysis can be performed by comparing differences in relative abundance of ^{32}P-labeled phosphoproteins, visualized by autoradiography, measuring relative intensities of the spot on the gel [10]. However these methods have the disadvantage to be time consuming and laborious and in the most cases they require large amounts of purified protein; moreover ^{32}P labelling could not be the method of choice for high-throughput proteome-wide analysis because of risks associated with handling radioactive compounds and because of the contamination of analytical instrumentation with radioactive materials. To avoid the use of radioactive probes, alternative approaches have been developed, based on the use of specific phospho-stains. Obviously they are less sensitive than radioactive probes, but they can be easily handled and used in high-throughput strategies. An example of these stains is the Pro-Q Diamond, a phospho-stain developed by Molecular Probes, specific for serine and threonine phosphorylation; Pro-Q Diamond is a fluorescent dye that can be used to stain phosphoproteins with sensitivity in the order of 1-20 ng of phosphoproteins as the lower limit. For a single protein the strength of the fluorescent signal can be correlated to the amount of phosphate groups. One of the most sensitive methodologies to analyse protein phosphorylation and to discriminate between serine, threonine and tyrosine phosphorylation, is based on the use of specific antibodies. Most commonly antibodies are used for phosphotyrosine

detection rather than phosphoserine/phosphothreonine because of their higher specificity and low cross-reactivity [11].
Recent efforts to eliminate the necessity to use radioactive probes or antibodies for the analysis of phosphorylation, have lead to the development of mass spectrometry based methods; these MS methods have emerged as the preferred tools for the analysis of proteins phosphorylation due to higher sensitivity, selectivity and rapidity than most biochemical techniques [12-13]. Phosphorylated proteins can be identified simply by the precise measurement by mass spectrometry of the molecular weight of the intact protein, just by looking at the increment of the protein mass compared with the unmodified one. However with this methodology the precise localization and nature of the modifying group. To better characterize the phosphorylated residues of a protein it is necessary to analyze the peptide mixture generated by enzymatic (usually trypsin) or chemical hydrolysis of the entire protein. With this approach phosphopeptides may be identified simply by compare experimental m/z values with theoretical ones, searching for mass increases of 80 Da. Also in this case it is possible to have ambiguities, which can be resolved by sequencing the peptide using tandem mass spectrometry [12,14]. A powerful mass spectrometric approach for phosphopeptides analysis [15-16] is based on the combined used of electrospray ionization and on-line liquid chromatography (LC-MS) [17].
However, the sub-stoichiometric nature of phosphorylation at a given site could be a problem also for MS analysis, thus resulting in sensitivity problems; moreover challenging fragmentation behaviour of many phosphopeptides under MS/MS conditions generally interfere with the elucidation of peptide sequences and with direct assignment of phosphorylation sites [18]. For these reasons separation technologies such as affinity, reverse phase, ion exchange chromatographies and capillary electrophoresis have been introduced prior to MS analysis to enrich the mixture in phosphopeptides or phosphoproteins that may otherwise be lost during mass spectrometric analysis. The enrichment step combined with the high sensitivity of MS technologies provides great potential for phosphoproteome characterization [19-23]. Recently, the use of miniaturized immobilized metal affinity chromatography (IMAC), in which phosphopeptides are bound non-covalently to solid supports that chelate Fe(III) or other trivalent metals has proved to be a potentially valuable method in phosphopeptides enrichment. With further refinement, this technique may offer the best performance for large-scale phosphorylation analysis [24-26]. Newly developed enrichment strategies are based on the use of MOAC protocols that take advantage of titania or zirconia as metal oxide chromatography modifiers [27]. Actually titanium dioxide (TiO_2) remains the most commonly used modifier.

IIII.1.2 Chemical tagging for the analysis of phosphorylation

Further several methods for selective enrichment of phosphopeptides use chemical modification of phosphorylated residues. In fact a series of chemical tagging procedures for phosphoproteomics, involving chemical reactions to facilitate the analysis of phosphorylated species, have been introduced [28], based mainly on β-elimination of phosphoserine (pSer) and phosphothreonine (pThr) residues under alkaline conditions, to form dehydroalanine (ΔSer) or β-methyldehydroalanine (ΔThr) respectively, and their ability to be targets for Michael type addiction with desired tags [29] (figure IIII.1). Treatment with an appropriate derivatizing agent has the potential to improve the detectability of phosphopeptides by mass spectrometric analysis. For example by adding a chemical compound containing a free sulfhydryl

group to β-eliminated residues, a new thiol group is created in the peptide that can be used as a linker for a molecular probe (for example a biotin probe) [19] or for direct affinity purification [30-31]. A smart approach was proposed by Knight and co-workers that used the combination of reaction and proteolysis to aid the assignment of phosphorylation sites using tandem MS analysis [32]. They converted pSer and pThr to aminoethylcysteine and β-methylaminoethylcysteine, respectively, both targets for lysine specific proteases. In this way peptides cleaved at the specific site of pSer or pThr were generated.
Via β-elimination Michael addition approach stable-isotope labelled tags can also be introduced. One limitation of β-elimination-based concepts is that pTyr residues cannot be addressed using this strategy. Furthermore is possible to have the contamination from peptides with other modifications that also undergo β-elimination, such as O-glycosylation.

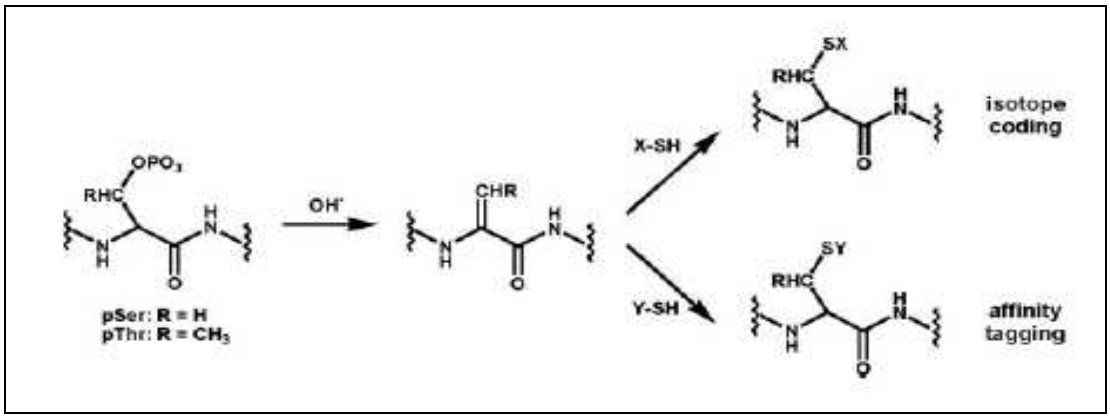


Figure IIII.1. *Tagging chemistry to address pSer and pThr residues.*

In this chapter, a chemical modification strategy of phosphorylation sites, based on β-elimination and Michael addition, is proposed aimed at the direct identification quantitative analysis of phosphorylated residues.

IIII.1.2 4000QTrap (QqLIT) for RIGhT strategy

RIGhT approaches are based on the selective MS analysis of species of interest by exploiting the fragmentation features of reagents chosen for selective labelling of residues of interest. As described in section II.1.2, iTRAQ reagents have the characteristics features of RIGhT reagents, since they have a reactive group toward the interesting residues, a linker group that is loss as neutral moiety during MS^2 experiment and a reporter group that leads to generate stable fragment ions selectively detectable in MS^2 experiment.
For the development of a RIGhT strategy that allows the identification and quantitation of phosphorylated proteins, iTRAQ was modified in order to obtain a donor for Michael type addiction. Although iTRAQ molecule was modified by cystamine addiction, its characteristic fragmentation pathway was not modified, thus it could still be used as RIGhT reagent. These characteristics fragmentation features were exploited by using an Applied Biosystems 4000Q-Trap mass spectrometer. By coupling the potential of this device and of the iTRAQ derivatives characteristics, it was possible to set up a high throughput LC-MS/MS method, based on *Precursor Ion Scan* (PIS) analysis, that allowed me to reduce the complexity of peptide mixtures by

mass spectrometric enrichment of phosphopeptides, rather than enriching and separating phosphopeptides prior to mass spectrometric analysis. Furthermore, the nanospray ESI source combined with nanoflow chromatographic system showed performances characterized by higher sensitivity and resolution, leading to the detection of samples even present at sub-stoichiometric amounts as phosphopeptides.
PIS was chosen as survey scan for its selectivity and its high scan rate. In this scan mode the Q_1 performs a scan across the full mass range; the ions separated are fragmented in the collision cell (q_2); finally the Q_3 is set to transmit only ions with the specific mass of the diagnostic fragment (m/z 114-117). Selected precursor ions are then submitted to a classical MS^2 experiment exploiting the LIT features of the QTrap. This MS based selection of labelled peptides substitutes others chromatographic steps so that the mass spectrometer can be considered as the second dimension of the RP-LC separation.

IIII.2 Materials and methods

Materials

Ammonium bicarbonate (AMBIC), guanidine, dithiothreitol (DTT), trypsin, iodoacetamide (IAM) and triethyl ammonium carbonate (TEAB) were purchased from Fluka. Bovine α-casein, tris(hydroxymethyl)aminomethane, tributylphosphine, cystamine, sodium dithionite as well as the MALDI matrix α-ciano-4-idrossicinnamic acid were purchased from Sigma. iTRAQ reagents were furnished by Applied Biosystems. Acetonitrile (ACN) was from Romil. Methanol and trifluoroacetic acid HPLC grade were purchased from Carlo Erba. All solvents were of the highest purity available from Baker. All other reagents and proteins were of the highest purity available from Sigma. Gel filtration columns PD-10 were from Pharmacia, whereas the pre-packed columns Sep-pak C-18 were from Waters.

iTRAQ-cysteamine synthesis

iTRAQ-cysteamine was obtained by reacting iTRAQ reagents (0.2 mmol/ml dissolved in acetonitrile) with cystamine (molar ratio 3:1). Reaction was carried out at room temperature for 1 hr, in 40% TEAB 500mM pH 8.0 and 60% ACN. The product was verified by MALDI-MS analysis. Then ACN was pushed away and the product was 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 using an Agilent Zorbax C18 column (4,6mm X 150mm) (Palo Alto, California) using a 10% to 65% linear gradient in 20 min from water to ACN and verified by MALDI-MS analysis.

Disulfide bridges reduction, Cys alkylation and proteolysis

Bovine α-casein was dissolved in denaturant buffer (urea 6M, Tris 300 mM pH 8.0, EDTA 10mM) and disulfide bridges were reduced with DTT (10-fold molar excess on the Cys residues) at 37 °C for 2 h and then alkylated by adding IAM (5-fold molar excess on thiol residues) at room temperature for 30 min in the dark. Protein samples were desalted by size exclusion chromatography on a Sephadex G-25M column (GE Healthcare). Fraction containing protein was lyophilized and the dissolved in 50 mM TEAB pH 8.0 buffer; trypsin digestion was performed using an enzyme/substrate ratio of 1/50 w/w at 37 °C for 16 h.

β-elimination reaction
β-elimination reactions were carried out by incubating the peptide mixture in 0.1 M $Ba(OH)_2$, at 37℃, for 90 min under nitrogen. Carbonic dioxide was then added. The precipitated barium carbonate was then removed by centrifugation at 13000 g/min for 10 min.

Michael addiction reaction
iTRAQ-cysteamine in Tris buffer 500mM pH 11.0 was added to the β-eliminated peptide mixture of α-casein. The mixture was allowed to react for 1 h in microwave oven [33] and then for 16h at 57℃.

Bacterial strains, growth conditions and protein extract preparation
Escherichia coli K12 strain was grown in aerobic conditions at 37℃ in LB medium. After 16 h, bacteria were harvested by centrifugation and resuspended in Buffer Z (25 mM HEPES pH 7.6, 50 mM KCl, 12.5 mM MgCl2, 1 mM DTT, 20% glycerol, 0.1% triton) containing 1 mM PMSF. Cells were disrupted by sonication. The suspension was centrifuged at 90000 x *g* for 30 min at 47℃. After centrifugation the protein concentration of the extract was determined with Bradford assay.

MALDI-TOF mass spectrometry
MALDI-TOF mass spectra were recorded using a Voyager DE-PRO mass spectrometer (Applied Biosystems, Framingham, USA). Prior to mass spectrometric analysis, peptide mixtures were purified using ZipTip pipette from Millipore (Billerica, MA, USA) using the recommended purification procedure. Peptides were eluted using 20 μL of 70% ACN, 0.1% HCOOH in water. A mixture of the eluted peptide solution and α-cyano-4-hydroxycinnamic acid (10 mg/ml in 70% ACN, 30% citric acid 50 mM in water) was loaded in 1/1 v/v ratio onto the metallic 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) as external standards. Raw data were analysed using Data Explorer software provided by Applied Biosystems and reported as monoisotopic masses.

nanoLC Mass Spectrometry
Peptide mixtures, obtained as previously described, were analysed by LC-MS using a 4000QTrap (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 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). Spectra acquisition was based on a survey precursor ion scan. Data were acquired and processed using Analyst software (Applied Biosystems).

Precursor Ion Scan analysis
Precursor ion scan analysis was performed over a mass range of m/z 400-1400 with a scan time of 2 s (with Q_1 set to low resolution and Q_3 set to unit resolution), with a nanospray voltage of 2 kV applied to an uncoated Silica Picotip (O.D. 150 μm, I.D. 10 μm, T.D. 15 μm) (New Objective, Woburn, MA). Precursors were fragmented in q_2

with a collision energy ramp of 30 to 80 V across the mass range. If a precursor of 114-117 m/z was detected above a pre-set threshold value (1500 cps), enhanced resolution scan was performed at 250amu/s to determine the charge state of the ion. The Information dependent acquisition (IDA) software automatically assigns a more accurate monoisotopic mass. Enhanced product ion (EPI) scans (MS/MS) were performed at 4000 amu/s and collision energies 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. The analysis was performed on the three most abundant ions and he entire cycle duration, including fill times and processing times, was less than 4 s.

IIII.3 Results and discussions

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, there is a demand to improve the methods for a reliable analysis of phosphopeptides. Phosphoproteins are usually present in low abundance in cells compared to their non-phosphorylated 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 of the analysis. Treatment with an appropriate derivatizing agent has the potential to improve the detectability of these analytes by MS.
Here it is 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^2 [34-35] analysis. A newly "in house" synthesized iTRAQ derivative is presented to selectively label and quantify phosphorylation sites. The procedure consists in the β-elimination of pSer/pThr residues and in the Michael addiction of the resulting α,β-unsatured residues with iTRAQ-cysteamine (figure IIII.3).

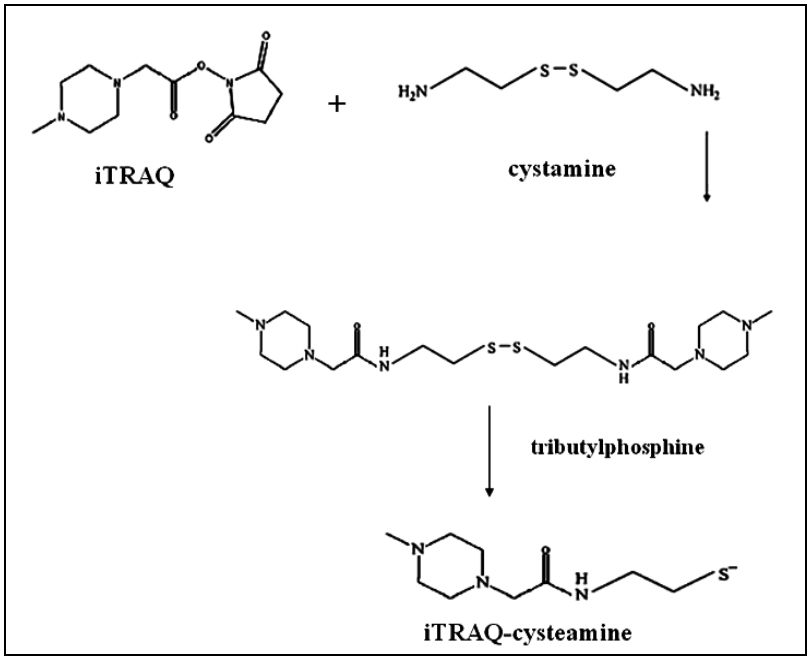


Figure IIII.3. *Reaction scheme for the synthesis of iTRAQ-cysteamine molecule.*

This reagent was chosen because of the need to have an SH moiety to perform Michael type addiction and to have an iTRAQ moiety to perform quantitative analysis. So the first objective was the optimization of the conditions for the synthesis of the iTRAQ-cystamine molecule and then for the Michael type addiction on β-eliminated phospho-residues. This derivatization introduces an iTRAQ moiety that fragments according to iTRAQ features (section I.2.5). Using the great capabilities of a QqLIT mass spectrometer, it was possible to take advantage of the distinctive m/z 114-117 fragments in MS^2. The methodology was first optimized by using a model phosphoprotein, bovine α-casein, and then to prove its wide enforceability this methodology applied to more complex samples.

IIII.3.1 Bovine α-casein

Bovine α-casein is a protein of bovine milk and it has eight phosphorylation sites, corresponding to serine residues 46, 48, 64, 66, 67, 68, 75 and 115; five of these residues are located on the same tryptic peptide, as shown by following sequence, where sites of phosphorylation are highlighted.

RPKHPIKHQGLPQEVLNENLLRFFVAPFPEVFGKEKVNELSKDIGSESTEDQAMEDIKQMEA
ESISSSEEIVPNSVEQKHIQKEDVPSERYLGYLEQLLRLKKYKVPQLEIVPNSAEERLHSMKE
GIHAQQKEPMIGVNQELAYFYPELFRQFYQLDAYPSGAWYYVPLGTQYTDAPSFSDIPNPIG
SENSEKTTMPLW

Commercial α-casein consists of α-casein S1 and α-casein S2 in traces and the preparations are usually contaminated with traces of β-casein.
A list of the theoretical tryptic phosphorylated peptides derived from α-caseins S1-S2 and their molecular masses, is shown in table IIII.1.

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

Table IIII.1. *Theoretical phosphorylated peptides originated by tryptic digestion of a mixture of α-caseins S1 and S2.*

IIII.3.2 Synthesis of iTRAQ-cysteamine

iTRAQ is known to react with ε-amino group of lysine as well as with N-terminal amino group. In this chapter, using chemical manipulations coupled with iTRAQ

labelling, a methodology capable of large-scale proteomic profiling and quantitative analysis of phosphorylation sites is shown. This method is based: i) on the selective modification of pSer/Thr residues with a novel iTRAQ derivative of cysteamine (figure IIII.3); ii) on the selective detection and identification of labelled peptides by exploiting the characteristic fragmentation pathway of iTRAQ-derivatives.
For synthesis of iTRAQ derivative of cystamine, it was used a slightly modified version of the protocol proposed by Applied Biosystems for iTRAQ labelling of peptides and proteins (see materials and methods section). iTRAQ-cystamine molecules obtained by reaction were analysed by MALDI-MS; as an example in figure IIII.4 it is reported a MALDI spectrum obtained from the analysis of an iTRAQ-cystamine sample; in this case it was chosen iTRAQ 115 as starting reagent. In the spectrum a signal at m/z 441.2 is evident that was attributed to the iTRAQ-cystamine molecule (theoretical molecular mass = 440.1 Da). Moreover it is present the signal at m/z 115.1 typical of the iTRAQ reagent chosen for the reaction.

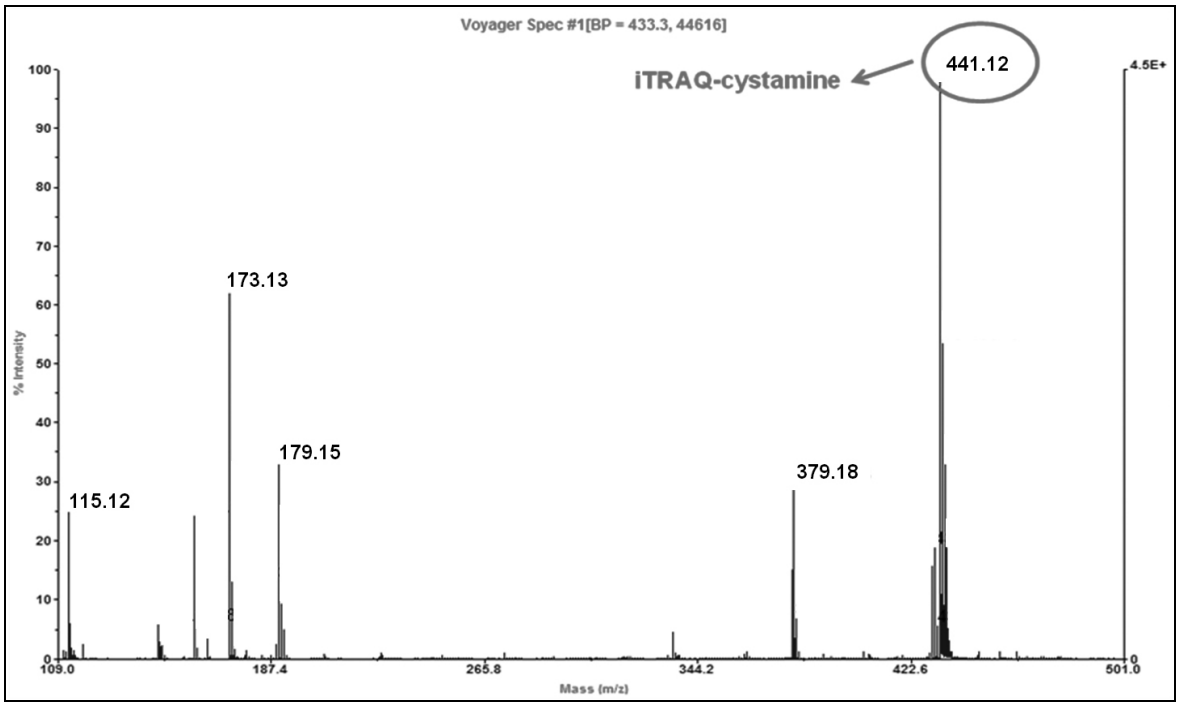


Figure IIII.4. *MALDI-MS spectrum of iTRAQ-cystamine molecule.*

For reduction of iTRAQ-cystamine to iTRAQ-cysteamine several reducing agents were used, in order to find the best reagent for my purposes; in particular the reduction was achieved by using DTT, tris(2-carboxyethyl)phosphine (TCEP) and tributylphosphine (TBP).
With regard to reducing power toward the cystamine disulfide bond, the best reducing agent was proven to be DTT while the worst one was TCEP; however DTT was difficult to remove from reaction mixture by RP-HPLC because it has physical-chemical characteristics similar to the iTRAQ-cysteamine molecule; moreover if some DTT was still present during Michael addiction, it could react as Michael donor, thus generating undesired by-products. For these reasons it was chosen to use TBP to perform iTRAQ-cystamine reduction; although it was not the best reagent in terms of reducing power, its hydrophobicity allowed me to easily separate it from iTRAQ-cysteamine molecules by RP-HPLC. After reduction, the product was purified by RP-

HPLC and the fractions collected were analysed by MALDI-MS. As expected, the reduction reaction generated a specie at m/z 222.1, as shown in figure IIII.5, corresponding to the reduction of disulfide bond of iTRAQ-cystamine

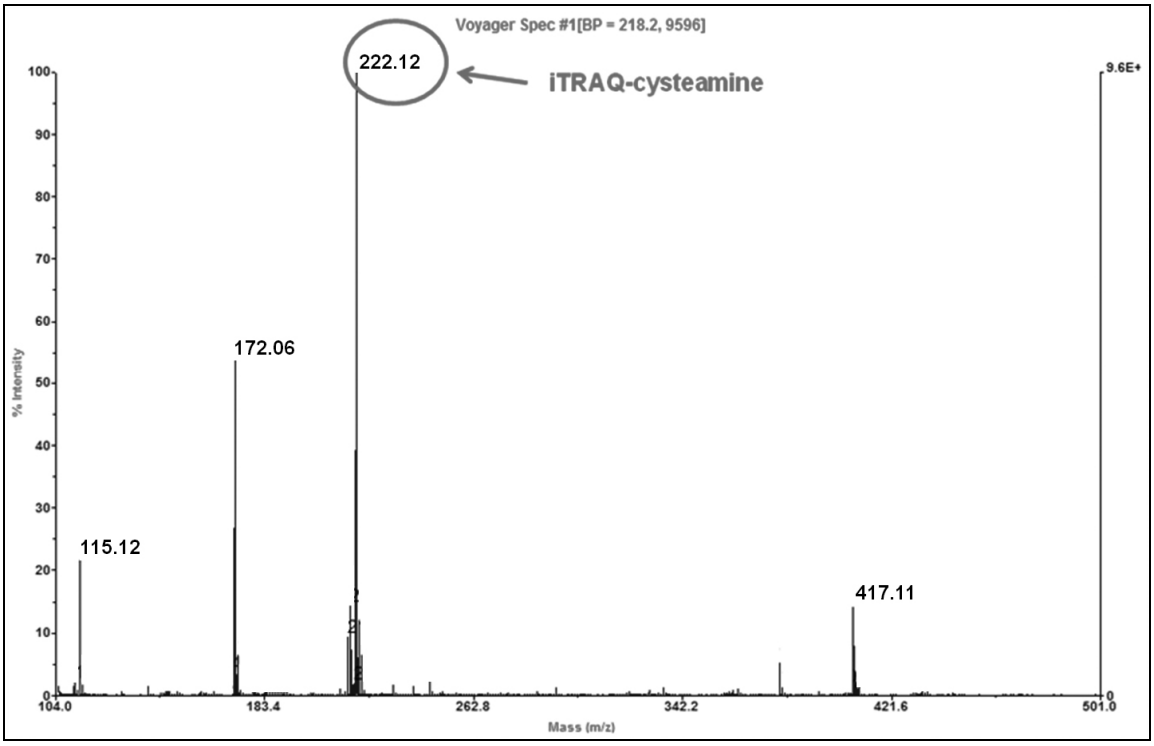


Figure IIII.5. *MALDI-MS spectrum of iTRAQ-cysteamine molecule.*

IIII.3.3 β-elimination reaction

Under strong alkaline conditions the phosphate moiety on pSer/pThr undergoes to β-elimination to form dehydroalanine (ΔSer) or β-methyldehydroalanine (ΔThr), respectively. The α,β-unsatured residues are strong Michael acceptors, which can readily react with a nucleophile [29,36-37].

The conditions for Michael type addition were optimized using bovine α-casein as model phosphoprotein. An aliquot of bovine α-casein was reduced with DTT and alkylated by IAM prior to tryptic digestion. The obtained peptide mixture was subjected to β-elimination reaction of phosphate moieties from pSer and pThr by using barium hydroxide as already reported [38]. The barium hydroxide was used replacing the previously reported NaOH [39-40] for its higher purity and its higher reactivity towards the pThr residues. To remove excess of reagent it was simply added solid carbon dioxide. The use of carbon dioxide permitted to remove the barium carbonate as a pellet on the bottom of a vial after centrifugation. Moreover when the pellet of barium carbonate was dissolved and analysed by MALDI-MS, no peptides were detected in the mass spectrum thus indicating the absence of a non-specific precipitation. The supernatant was equally analysed by MALDI mass spectrometry, using a Voyager-DE STR mass spectrometer. Prior to MS analysis, peptide mixtures were purified using ZipTip pipette, using the recommended purification procedure. As expected the MALDI spectrum (figure IIII.6) showed the occurrence of signals not attributable to tryptic peptides of phosphorylated bovine α-casein; these signals showed a ΔM of -98 Da respect to phosphorylated peptides,

and were attributed to corresponding β-eliminated peptides. As an example, the ions at m/z 1854.6 and m/z 1563.5 were assigned, respectively, to peptide 104-119 and 106-119 of α-casein, occurring 98 Da lower than that expected, indicating that the pSer115 was converted in dehydroalanyl residue after β-elimination reaction. The yield of reaction was almost 100%.

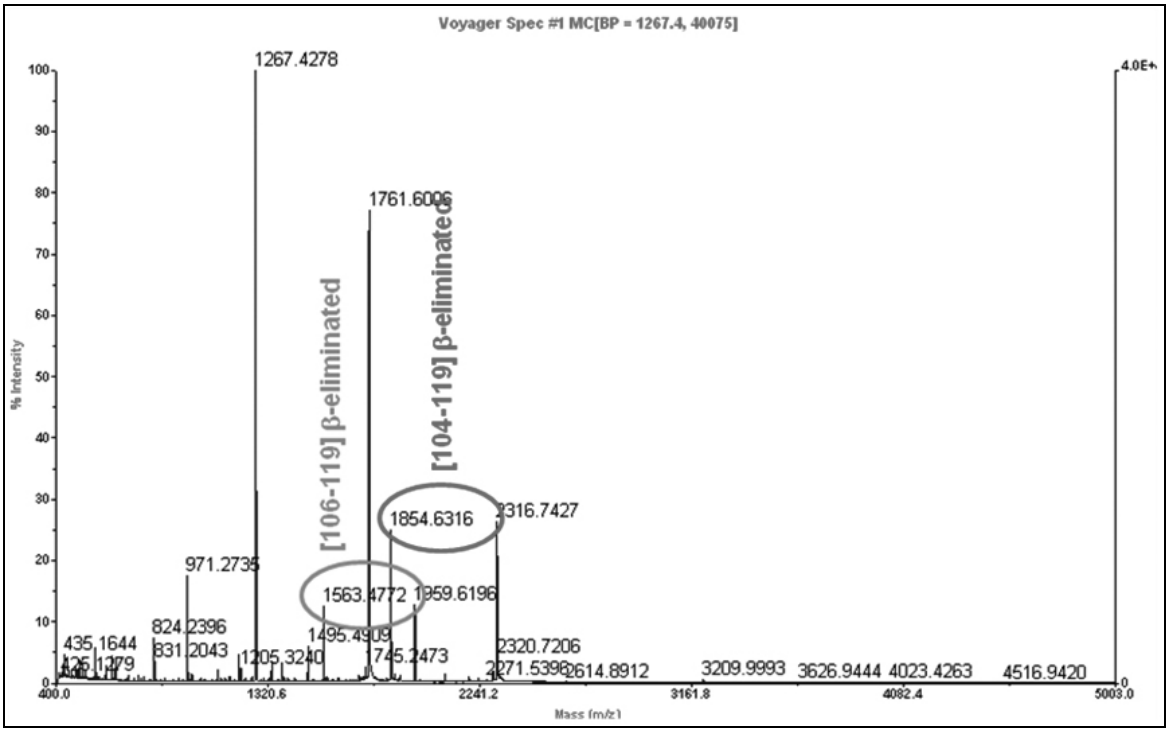


Figure IIII.6. *MALDI-MS spectrum of peptide mixture from α-casein following β-elimination reaction. As evident two signals are present at m/z 1563.48 and 1854.63 attributable to newly β-eliminated peptides occurring at minus 98 Da respect to phosphorylated forms.*

The α,β-unsatured residues are Michael acceptor, which can readily react with a nucleophile.

IIII.3.4 iTRAQ-cysteamine addiction

The β-eliminated peptide mixture was modified with iTRAQ-cysteamine via Michael type addition by following the procedure described in materials and methods section. The rationale behind the choice of this reagent was due to the possibility to use the iTRAQ moiety for selective mass spectrometric analysis through precursor ion scan mode and at the same time to use it for quantitative analysis of phosphopeptides. The addition reaction resulted in the replacement of what was formerly a phosphate moiety with an iTRAQ moiety. The interpretation of MALDI spectrum obtained from the analysis of a labelled peptide mixture (figure IIII.7) revealed the presence of two signals at m/z 1784.8 and m/z 2075.4, attributed to β-eliminated peptides 106-119 and 104-119, respectively, modified by iTRAQ-SH (ΔM = + 221 Da).

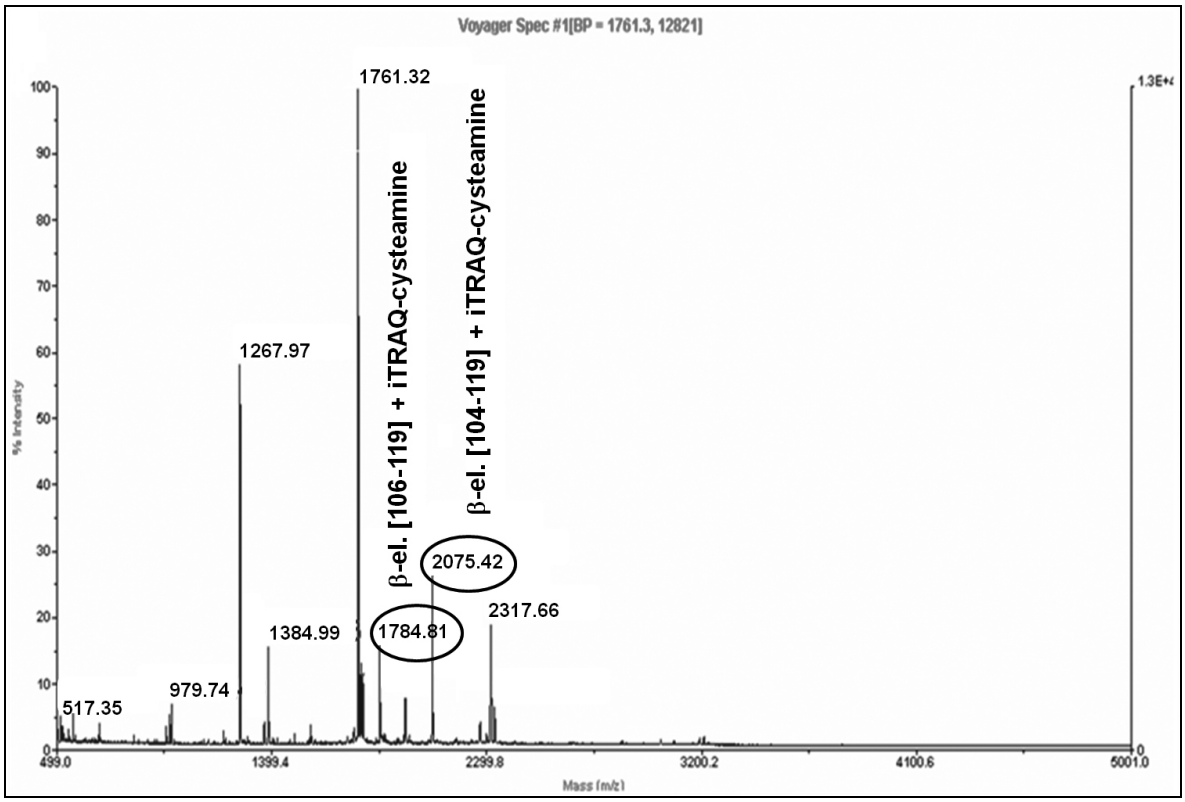


Figure IIII.7. *MALDI-MS spectrum of peptide mixture from α-casein following Michael type addiction with iTRAQ-cysteamine molecule. As evident two signals are present at m/z 1784.81 and 2075.43 attributable to the β-eliminated peptides 106-119 and 104-119, respectively, modified by iTRAQ-cysteamine (ΔM = + 221 Da).*

The relative amounts of iTRAQ-modified peptide peaks and β-eliminated peptide peaks in MALDI spectra indicates that the yield of the addition reaction was quite good but not quantitative (about 70%), as expected for a typical Michael addition reaction [31]. However, the direct MALDI-MS analysis of the entire peptide mixture suffered from suppression phenomena leading to the disappearance of other species; in fact it was not possible to directly identify other β-eliminated and iTRAQ-cysteamine modified peptides both from α-casein S1 and S2, despite of the fact that it should be other phosphorylated residues in both proteins. To solve this problem, peptide mixtures were analysed by LC-MS/MS exploiting the characteristic features of a 4000QTrap system.

IIII.3.5 Analysis by LC-MS/MS

An aliquot of the modified peptide mixture was submitted to LC-MS/MS analysis by using 4000QTrap coupled to an 1100 nano HPLC system. We settled an experiment using as survey scan a *precursor ion scan* (PIS) for the iTRAQ reporter ions at m/z 117/114. The PIS is a typical triple quadrupole scan mode which can be used when an ion loses a diagnostic charged fragment during MS^2 experiments. The use of PIS as survey scan allowed to enrich the mixture in modified peptides during MS analysis; the reconstructed ion chromatogram for the *precursor ion scan* mode of α-casein peptide mixture after Michael type addiction is reported in figure IIII.8.

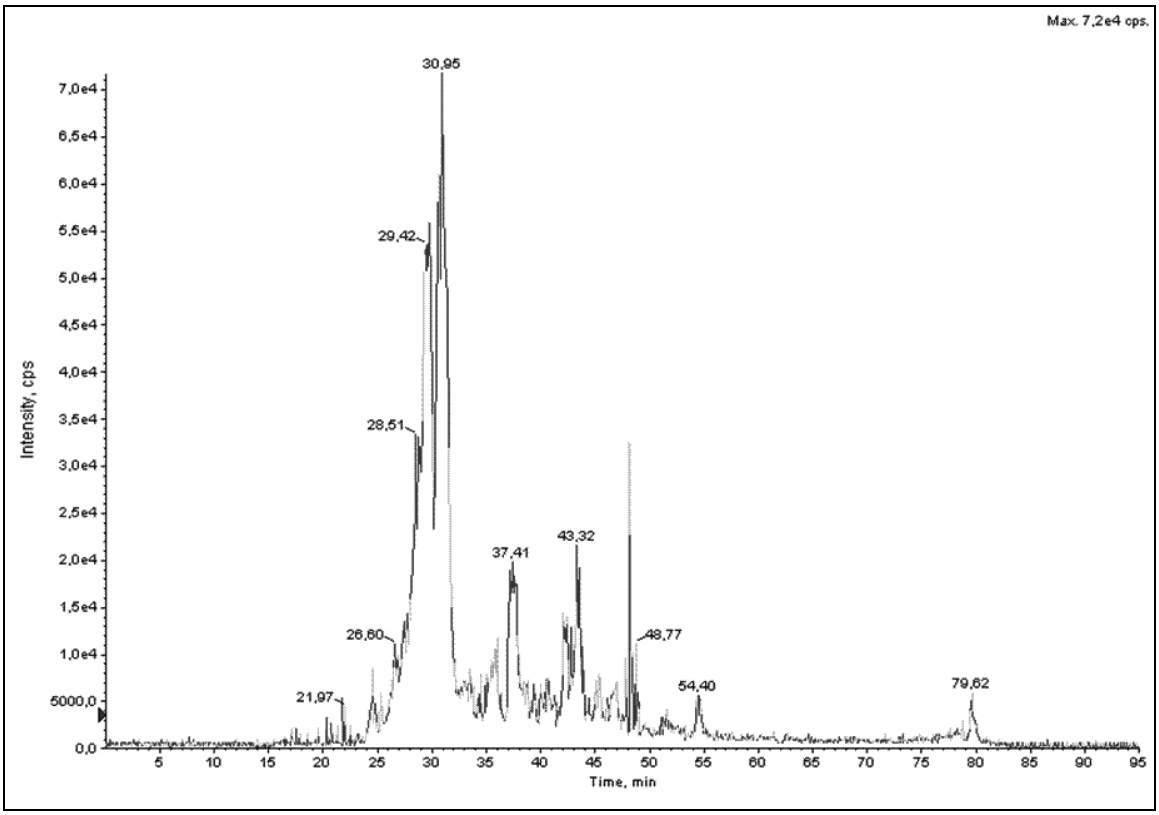


Figure IIII.8. *Total Ion Current of Precursor Ion Scan analysis.*

This chromatogram showed a simple profile not attributable to the analysis of a complex peptide mixture; this is due to the fact that only peptides generating iTRAQ reporter ions during MS^2 analysis could be revealed by the mass spectrometer. MS^2 spectra of the species eluted from 28.5 to 30.9 min were attributed to the β-eliminated phosphopeptides 104-119 and 106-119 of α-casein carrying an iTRAQ-cysteamine moiety. As an example, the MS/MS spectrum of the modified peptide 106-119 is reported in figure IIII.9. This ion was stable during collision induced dissociation and provided easily interpretable product ion spectra. In fact, the y and b fragment ions still retained the modifying group on the β-eliminated Ser residue, thus allowing the exact localisation of the phosphorylation site.

It is worth considering that α-casein contains also a di-phosphorylated and a penta-phosphorylated tryptic peptides which escaped to be detected using this approach probably due to the chemical-physical characteristics of the modified fragments. In fact the introduction of a high number of iTRAQ moieties increases the hydrophobicity of peptides; moreover the low efficiency of Michael addition results dramatically increased when more than one sites are present in the same peptide.

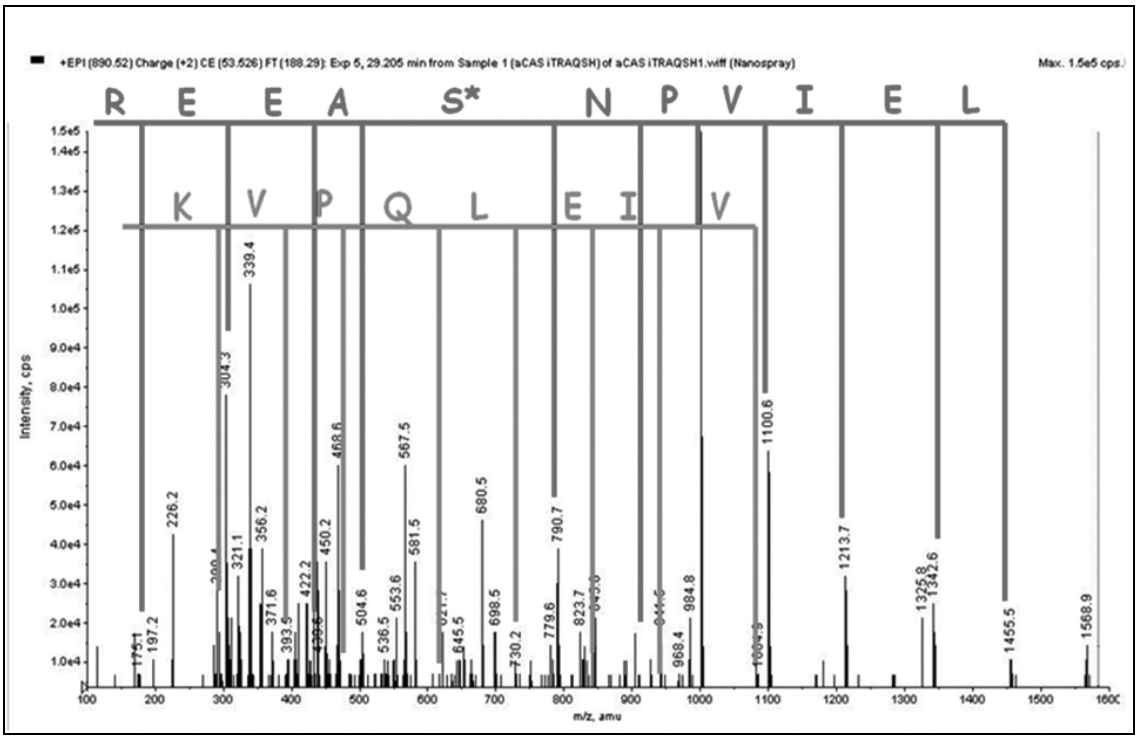


Figure IIII.9. *MS/MS spectrum of peptide 104-119 modified with iTRAQ-cysteamine on Ser 115; the y and the b series are reported, highlighting the presence of the modified residue.*

IIII.3.6 Analysis of a complex mixture

In a proof of principle experiment, in order to demonstrate the ability of this analytical method to specifically locate phosphorylated residues in a proteomics analysis, 10 mg of an entire cellular extract of *Escherichia coli* cells, was spiked with 5 μg of β-eliminated α-casein tryptic mixture. This complex peptide mixture was then modified by reaction with iTRAQ-cysteamine via Michael-type addition by following the procedure described above. 1 pmol of the peptide mixture was then submitted to LC-MS/MS analysis using precursor ion scan mode, as described above. Figure IIII.10 shows the reconstructed ion chromatogram for the selective iTRAQ fragmentation in MS^2 mode.
Data obtained by MS/MS analysis in precursor ion scan mode were exploited both to confirm the presence of modified α-casein in the mixture and to identify modified peptides By taking advantage of the flexibility of the in house Mascot software. The software was properly modified by adding in the "modification file" the value of ΔM + 221 Da for iTRAQ labelled β-eliminated pSer and pThr residues. For Mascot analysis it was necessary to set the software supplying key information, e.g. proteolytic enzyme, fixed modifications, variable modifications, sample taxonomy, eventually missed cleavages, etc. Mascot search results are reported in figure IIII.11.

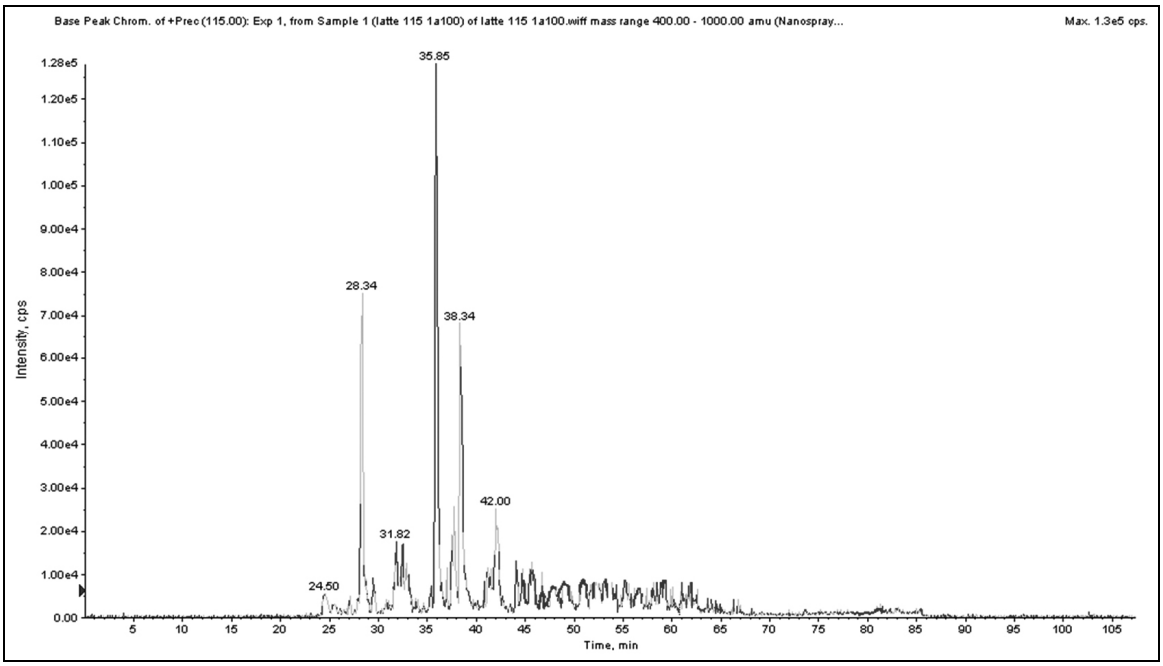


Figure IIII.10. *Total Ion Current of Precursor Ion Scan analysis obtained from the analysis of a complex peptide mixture after iTRAQ-cysteamine modification of β-eliminated phosphopeptides of α-casein. Despite the complexity of the mixture the chromatogram appeared to be very simple, due to the PIS simplification of the sample in analysis; in fact the only peaks presents in the chromatograms are attributable to iTRAQ modified phosphopeptides.*

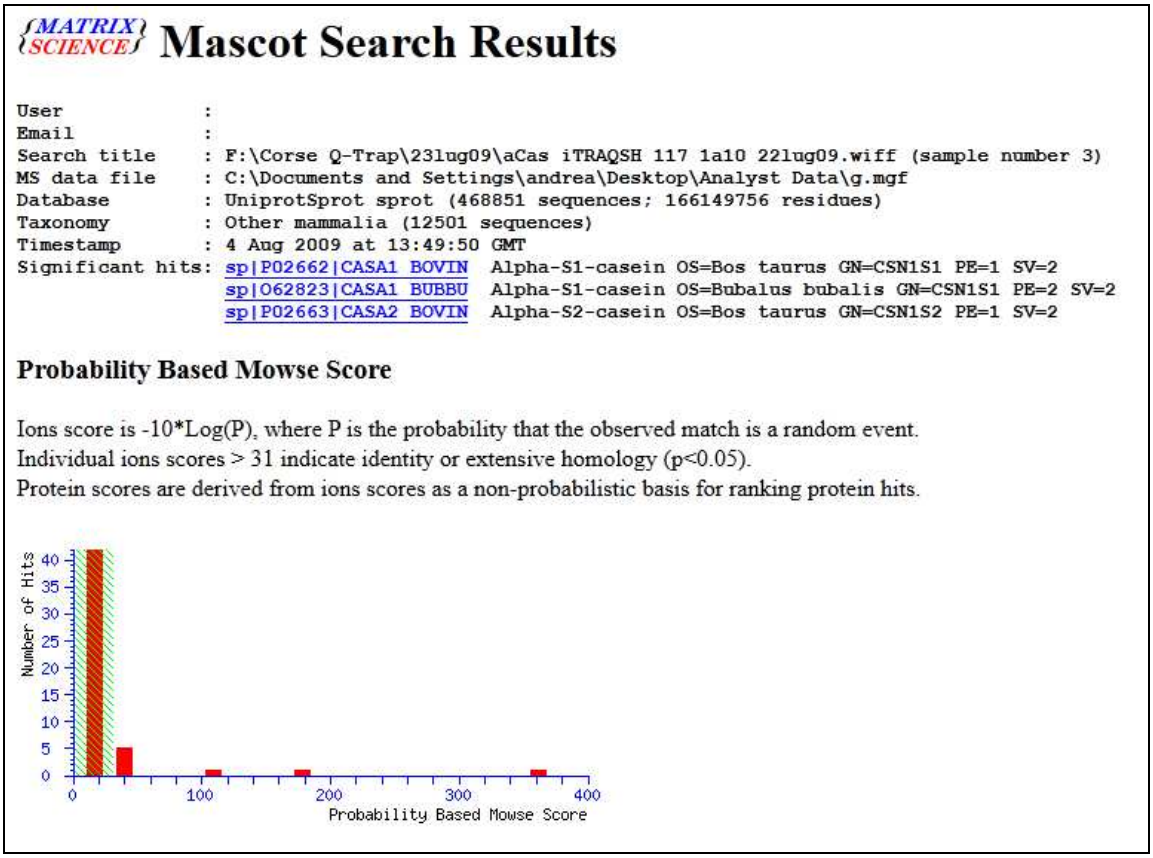


Figure IIII.11. *Mascot search result of the data form LC-MS/MS analysis of the complex protein mixture.*

As shown by Mascot data, this analytical strategy led to identification of only α-casein S1 and α-casein S2, even if a more complex mixture was analysed by LC-MS/MS. Moreover exploiting the fragmentation features of iTRAQ molecules in combination with PIS analysis, it was possible to selectively isolate from the entire mixture only modified peptides. A summary of peptides identified during this analysis is reported in table IIII.3. However, other peptides appear in the table not related to modified phosphopeptides. This effect could be due to the presence of different fragment ions in the region 113-114 or 116-117 in combination with the ion resolution of quadrupoles during PIS analysis; in fact in order to increase the sensitivity of mass spectrometric analysis, Q_1 and Q_3 peak widths were set to 0.7 amu.
It is interesting to note the presence of two modified peptides that were not present in previous analysis. In fact it was possible to identify the di-phosphorylated peptide of α-casein S1, with both phosphorylated residues (serine 46 and serine 48) labelled with an iTRAQ-cysteamine moiety, and a modified phosphopeptide of α-casein S2. These data prove the selectivity and the sensitivity of the strategy proposed; indeed, despite the increase in sample complexity, it was possible to detect modified peptides.

Protein	Sequence	Monoisotopic mass	Modifications
α-casein S1	YLGYLEQLLR	1267.80	
α-casein S1	FFVAPFPEVFGK	1384.93	
α-casein S1	HQGLPQEVLNENLLR	1760.19	
α-casein S1	YKVPQLEIVPNSAEER	2075.36	iTRAQ-cysteamine
α-casein S1	VPQLEIVPNSAEER	1784.72	iTRAQ-cysteamine
α-casein S1	DIGSESTEDQAMEDIK	2174.74	2 iTRAQ-cysteamine
α-casein S2	FALPQYLK	979.59	
α-casein S2	ALNEINQFYQK	1366.99	
α-casein S2	TVDMESTEVFTK	1589.94	iTRAQ-cysteamine
α-casein S2	TVDMESTEVFTKK	1718.37	iTRAQ-cysteamine
α-casein S2	KTVDMESTEVFTKK	1846.35	iTRAQ-cysteamine

Table IIII.3. *Summary of peptides identified during LC-MS/MS analysis of the complex protein mixture. Despite the complexity of the sample it was possible to identify phosphorylated peptides of both α-casein S1 and S2.*

IIII.3.7 Quantitative analysis of protein phosphorylation

To address the problem of quantitation, having added an iTRAQ moiety to pSer/pThr residues, it is possible to follow basically the strategy outlined by Gygi *et al.* [41]. The peptides were quantified by measuring, in the mixture, the relative signal intensities for pairs of peptide ions of identical sequence, differentially labelled with two different forms of iTRAQ. The mass difference within the iTRAQ reporter ions reliably allows a separation between such ions; the ratio between the peak areas of reporter ions

provide an accurate measurement of the relative abundance of labeled peptides and hence of the related proteins in the original samples. In fact the 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 phosphorylation of a peptide from two different samples, iTRAQ-cysteamine 115 and 116 were used to label samples containing stoichiometric concentrations of α-casein in ratios of 1:1, 1:2 and 1:4. The peptide mixture were analysed by LC-MS/MS analysis by using a *precursor ion scan* (PIS) for the reporter ions at m/z 115/116. It was already demonstrated by a previous work [42] that the choice of the fragment for PIS analysis is un-relevant, due to the isobaric characteristic of the iTRAQ derivatives. A MS/MS spectrum obtained by PIS analysis of a modified peptide mixture is reported in figure IIII.12. The signals present in the spectrum led to the reconstruction of almost the entire sequence of the β-eliminated phosphopeptide 104-119 of α-casein S1 carrying an iTRAQ-cysteamine moiety. Moreover in the spectrum it is evident the presence of iTRAQ reporter ions in the low mass region that were used to perform relative quantitative analysis of the mixtures; by using the Analyst software tool for calculation of peak areas (Applied Biosystems) it was possible to compare the areas of signals relative to reporter ions at m/z 115 and 116.

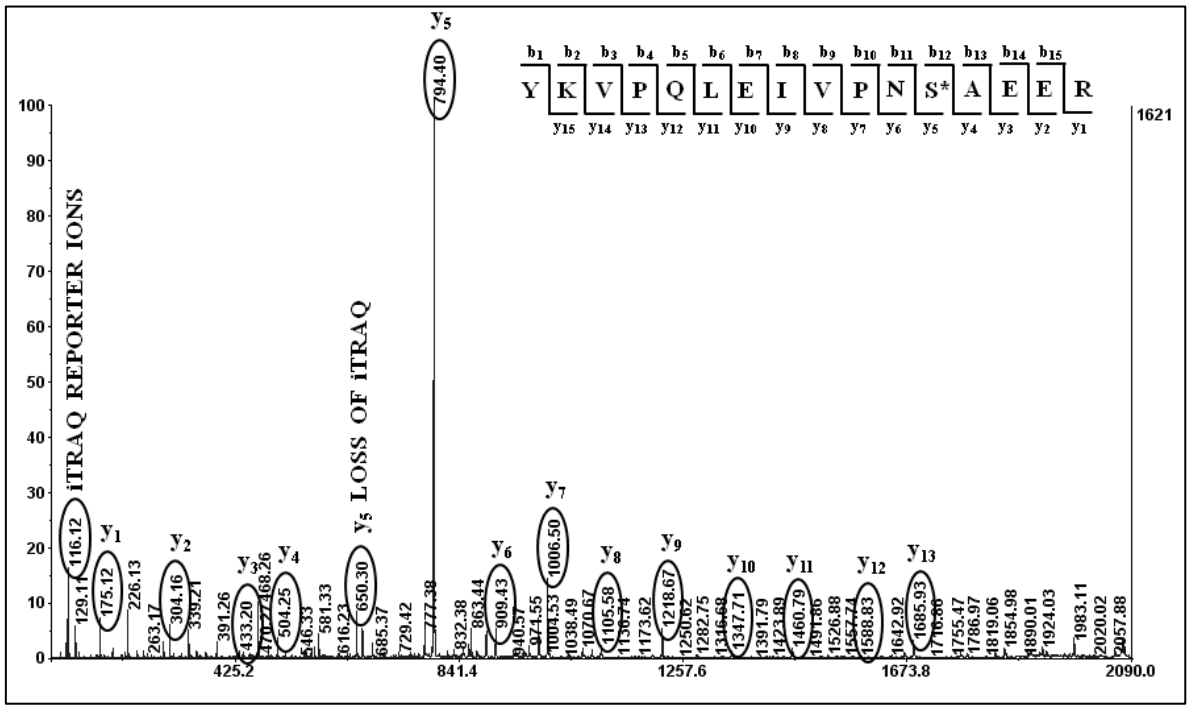
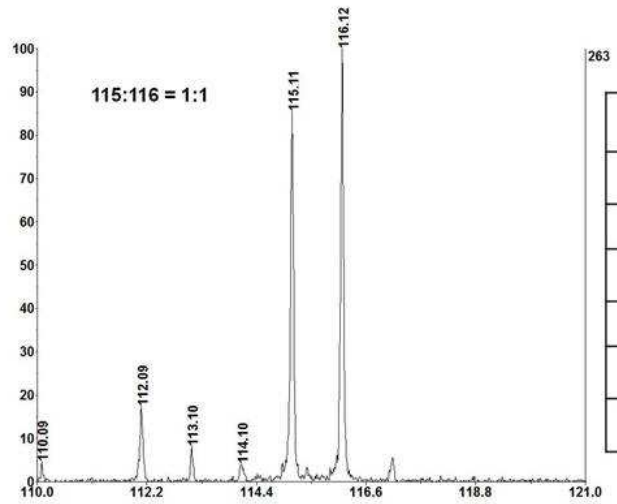
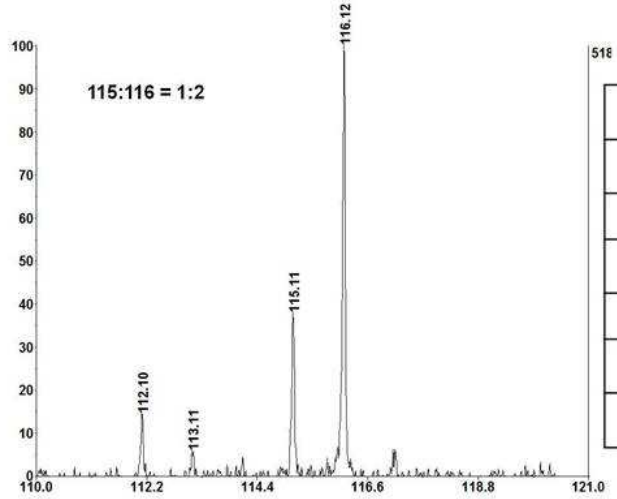


Figure IIII.12. *MS/MS spectrum of the of the β-eliminated phosphopeptide 104-119 carrying an iTRAQ-cysteamine moiety. In the spectrum it is evident the presence of iTRAQ reporter ions that can be used to perform relative quantitative analysis.*

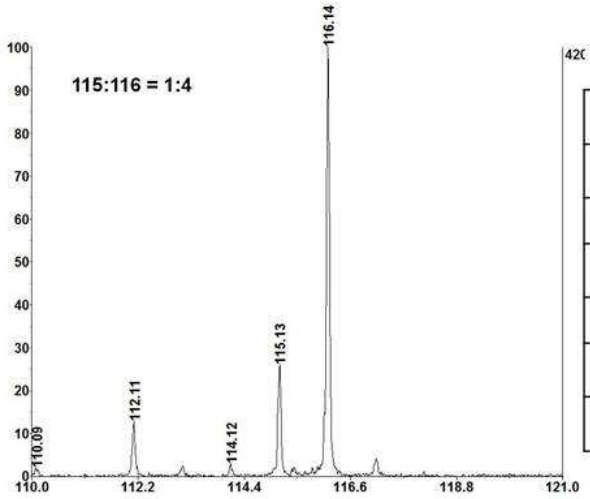
Magnified sections of MS/MS spectra obtained from the analysis of the three different mixtures, showing the region of iTRAQ reporter ions, are reported in figure IIII.13 in three different panels.



Experiment #	m/z	Peak area (counts)
1	115.1	2206,36
	116.1	2464,91
2	115.1	2516,19
	116.1	2444,42
3	115.1	2497,51
	116.1	2449,45



Experiment #	m/z	Peak area (counts)
1	115.1	2001,99
	116.1	4233,24
2	115.1	2141,84
	116.1	4278,22
3	115.1	2009,84
	116.1	4197,23



Experiment #	m/z	Peak area (counts)
1	115.1	1085,53
	116.1	4390,65
2	115.1	1127,91
	116.1	4380,24
3	115.1	2009,84
	116.1	4482,32

Figure IIII.13. *Magnified sections of MS/MS spectra, obtained from the analysis of differentially labeled α-casein peptide mixtures, showing the region of iTRAQ reporter ions and the peak areas used for relative quantification.*

Peak areas of the iTRAQ reporter ions are reported in table IIII.4; analyses were performed in triplicate. Experimental ratios resulted in excellent agreement with the theoretical ratios (table IIII.4 and figure IIII.13).

116/115 Theoretical ratio	116/115 Experimental ratios
1	1,02 ± 0,08
2	2,06 ± 0,06
4	3,99 ± 0,05

Table IIII.4. *Summary of the results obtained by quantitative analysis of the tree differentially labelled aliquots of α-casein peptide mixtures.*

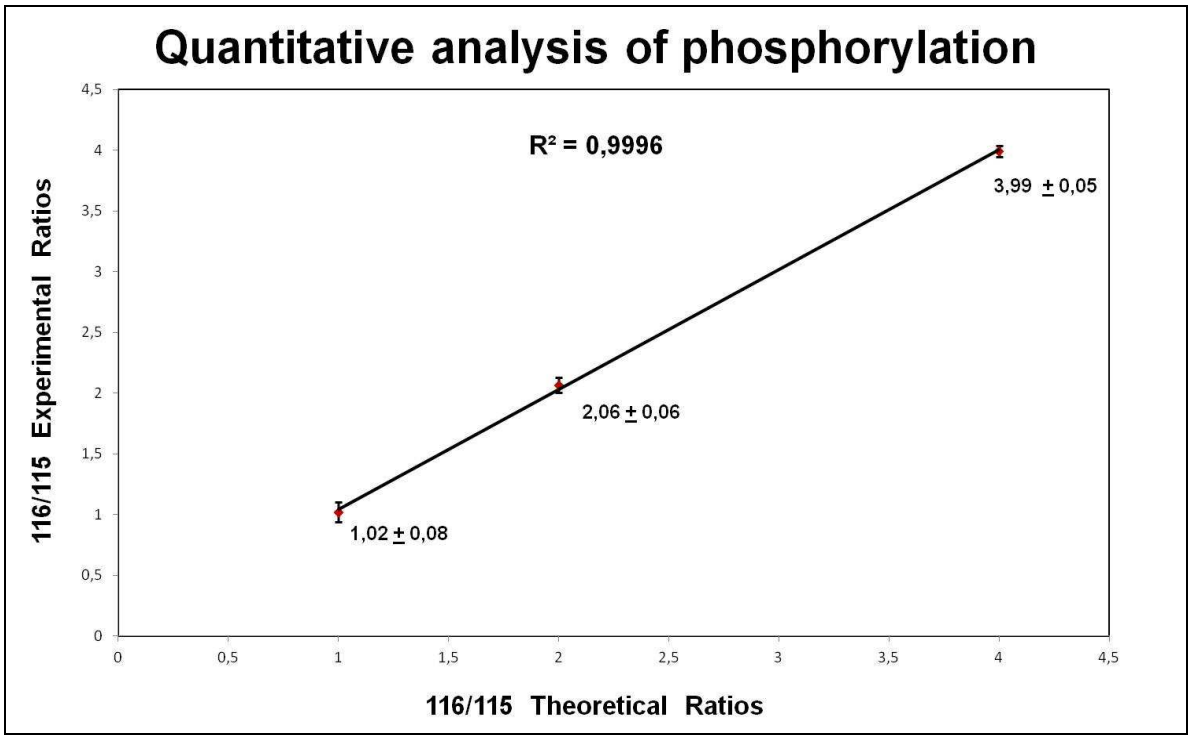


Figure IIII.14. *Graphical representation of the results obtained by quantitative analysis of the tree differentially labelled aliquots of α-casein peptide mixture. Experimental ratios are in good agreement with theoretical ratios.*

IIII.4 Conclusions

This chapter reports an integrated simple methodology for the qualitative and quantitative analysis of protein phosphorylation, which involves chemical replacement of the phosphate moieties by iTRAQ moieties. This methodology can be classified as RIGhT procedure, since it makes use of a class of specific reagents and takes advantage of novel mass spectrometer such as a liner ion trap. Moreover it was proven to lead to the simultaneous localisation and quantification of phosphorylation sites both in model proteins and in more complex samples.
The 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. The quantitative analysis can be performed by using the different isotopically labelled forms of iTRAQ reagents to selective label the modified sites, thus leading to a quantitative description of phosphorylation states.
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. Here the strategy proposed, taking in account the availability of linear trap to select specific labelled peptides giving rise to diagnostic MS^2 product ions, resulted to be of useful in the phosphoproteins analysis for proteomics purpose.

IIII.5 References

1. Ross, P.L.; Huang, Y.N.; Marchese, J.N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; Feng, He; Jacobson, A.; Pappin, D.J. *Mol Cell Proteom*. **2004**, *3*, 1154-1169.
2. Clifton, J.G.; Huang, F.; Kovac, S.; Yang, X.; Hixson, D.C.; Josic, D. *Electrophoresis*. **2009**, *30* (20), 3636-3646.
3. Hakimov, H.A.; Walters, S.; Wright, T.C.; Meidinger, R.G.; Verschoor, C.P.; Gadish, M.; Chiu, D.K.; Strömvik, M.V.; Forsberg, C.W.; Golovan, S.P. *Proteomics*. **2009**, *9* (16), 4000-4016.
4. Zamò, A.; Cecconi, D. *J Proteomics*. **2010**, *73* (3), 508-520.
5. Ow, S.Y.; Salim, M.; Noirel, J.; Evans, C.; Rehman, I.; Wright, P.C. *J Proteome Res*. **2009**, *8* (11), 5347-5355.
6. Yan, J.X.; Packer, N.H.; Gooley, A.A.; Williams, K.L. *J Chromatogr A*. **1998**, *808*, 23-41.
7. Schroeder, M.J.; Shabanowitz, J.; Schwartz, J.C.; Hunt, D.F.; Coon, J.J. *Anal Chem*. **2004**, *76*, 3590-3598.
8. Kaufmann, H.; Bailey, J.E.; Fussenegger, M. *Proteomics*. **2001**, *1* (2), 194-199.
9. Bendt, A.K.; Burkovski, A.; Schaffer, S.; Bott, M.; Farwick, M.; Hermann, T. *Proteomics*. **2003**, *3* (8), 1637-1646.
10. Cohen, P. *Trends Biochem Sci*. **2000**, *25* (12), 596-601.
11. McLachlin, D.T.; Chait, B.T. *Curr Opin Chem Biol*. **2001**, *5* (5), 591-602.
12. Pandey, A.; Andersen, J.S.; Mann, M. *Sci STKE*. **2000**, *2000* (37), PL1.
13. Janek, K.; Wenschuh, H.; Bienert, M.; Krause, E. *Rapid Commun Mass Spectrom*. **2001**, *15* (17), 1593-1599.
14. Biemann, K. *Methods Enzymol*. **1990**, *193*, 455-479.
15. Poulter, L.; Ang, S.G.; Gibson, B.W.; Williams, D.H.; Holmes, C.F.; Caudwell, F.B.; Pitcher, J.; Cohen, P. *Eur J Biochem*. **1988**, *175* (3), 497-510.
16. Ohguro, H.; Palczewski, K.; Ericsson, L.H.; Walsh, K.A.; Johnson, R.S. *Biochemistry*. **1993**, *32* (21), 5718-5724.

17. Ballif, B.A.; Villen, J.; Beausoleil, S.A.; Schwartz, D., Gygi, S.P. *Mol Cell Proteomics*. **2004**, *3* (11), 1093-1101.
18. Gafken, P.R.; Lampe, P.D. *Cell Commun Adhes*. **2006**, *13* (5-6), 249-262.
19. Oda, Y., Nagasu, T.; Chait, T. *Nat Biotechnol*. **2001**, *19*, 379-382.
20. Pinkse, M.W.; Uitto, P.M.; Hilhorst, M.J.; Ooms, B.; Heck, A.J. *Anal Chem*. **2004**, *76* (14), 3935-3943.
21. Larsen, M.R.; Thingholm, T.E.; Jensen, O.N.; Roepstorff, P.; Jørgensen, T.J. *Mol Cell Proteomics*. **2005**, *4* (7), 873-886.
22. Chen, G.; Pramanik, B.N. *Drug Discov Today*. **2009**, *14* (9-10), 465-471.
23. Gafken, P.R. *Methods Mol Biol*. **2009**, *527*, 159-172.
24. Thingholm, T.E.; Jensen, O.N.; Larsen, M.R. *Methods Mol Biol*. **2009**, *527*, 67-78.
25. Thingholm, T.E.; Jensen, O.N.; Larsen, M.R. *Methods Mol Biol*. **2009**, *527*, 47-56.
26. Sun, X.; Chiu, J.F.; He, Q.Y. *Expert Rev Proteomics*. **2005**, *2* (5), 649-657.
27. Sugiyama, N.; Masuda, T.; Shinoda, K.; Nakamura, A.; Tomita, M.; Ishihama, Y. *Mol Cell Proteomics*. **2007**, *6*, 1103-1109.
28. Leitner, A.; Lindner, W. *Methods Mol Biol*. **2009**, *527*, 229-243.
29. Chen, M.; Su, X.; Yang, J.; Jenkins, C.M.; Cedars, A.M.; Gross, R.W. *Anal Chem*. **2010**, *82* (1), 163-171.
30. McLachlin, D.T.; Chait, B.T. *Anal Chem*. **2003**, *75*, 6826-6836.
31. Amoresano, A.; Marino, G.; Cirulli, C.; Quemeneur, E. *Eur J Mass Spectrom*. **2004**, *10* (3), 401-412.
32. Knight, Z.A.; Schilling, B.; Row, R.H.; Kenski, D.M. *Nat Biotechnol*. **2003**, *21*, 1047-1054.
33. Tzeng, Y.K.; Chang, C.C.; Huang, C.N.; Wu, C.C.; Han, C.C.; Chang, H.C. *Anal Chem*. **2008**, *80* (17), 6809-6814.
34. Olsen, J.V.; Mann, M. *Proc Natl Acad Sci USA*. **2004**, *101*, 13417-13422.
35. Venable, J.D.; Dong, M.Q.; Wohlschlegel, J.; Dillin, A.; Yates, J.R. *Nature Methods*. **2004**, *1*, 1.
36. Conrads, T.P.; Issaq, H.J.; Veenstra, T.D. *Biochem Biophys Res Commun*. **2002**, *290* (3), 885-890.
37. Kim, J.S.; Kim, J.; Oh, J.M.; Kim, H.J. *Anal Biochem*. **2011**, *414* (2), 294-296.
38. Amoresano, A.; Monti, G.; Cirulli, C.; Marino, G. *Rapid Commun Mass Spectrom*. **2006**, *20* (9), 1400-1404.
39. Li, W.; Boykins, R.A.; Backlund, P.S.; Wang, G.; Chen, H.C. *Anal Chem*. **2002**, *74* (22), 5701-5710.
40. Mega, T.; Nakamura, N.; Ikenaka, T. *J Biochem*. **1990**, *107*, 68-72.
41. Gygi, S.P.; Rist, B.; Gerber, S.A.; Turecek, F.; Gelb, M.H., Aebersold, R. *Nature Biotechnol*. **1999**, *17*, 994-999.
42. Chiappetta, G.; Corbo, C.; Palmese, A.; Galli, F.; Piroddi, M.; Marino, G.; Amoresano, A. *Proteomics*. **2009**, *9* (6), 1524-1537.

Table of abbreviations

ACN	**Acetonitrile**
AMBIC	**ammonium hydrogen carbonate**
α-ciano	**alpha-cyano-4-hydroxycinnamic acid**
BSA	**bovine serum albumin**
CID	**collision induced dissociation**
DC	**direct current**
DNS	**dansyl**
DNSH	**dansyl-hydrazide**
DTT	**dithiothreitol**
ESI	**Electrospray Ionisation**
IAM	**iodoacetamide**
ICAT	**Isotope Coded Affinity Tag**
IT	**ion trap**
iTRAQ	**Isobaric Tag for Relative and Absolute Quantification**
iTRAQH	**iTRAQ-hydrazide**
LC	**liquid chromatography**
LIT	**linear ion trap**
m/z	**mass-to-charge ratio**
MALDI	**Matrix-Assisted Laser Desorption/Ionisation**
MS	**mass spectrometry**
N-BSA	**nitrated bovine serum albumin**
NH_2-Tyr	**3-aminotyrosine**
NO_2-Tyr	**3-nitrotyrosine**
PIS	**Precursor Ion Scan**
PTMs	**post-translational modifications**
Q	**quadrupole**
QqQ	**triple quadrupole**
RF	**radio frequency**
RIGhT	**Reporter Ion Generating Tag**
RNasi A	**bovine ribonuclease A**
RNS	**reactive nitrogen species**
ROS	**reactive oxygen species**

RP-HPLC	**Reverse Phase High Performance Liquid Chromatography**
SDS	**Sodium Dodecyl Sulphate**
S/N	**signal to noise ratio**
TOF	**Time Of Flight**

Pubblications

1) Chiappetta G, Corbo C, **Palmese A**, Galli F, Piroddi M, Marino G, Amoresano A.
   Quantitative identification of protein nitration sites.
   Proteomics. 2009 Mar;9(6):1524-37. Erratum in: Proteomics. 2009 Jun;9(11):3220.

2) Amoresano A, Carpentieri A, Giangrande C, **Palmese A**, Chiappetta G, Marino G, Pucci P.
   Technical advances in proteomics mass spectrometry: identification of post-translational modifications.
   Clin Chem Lab Med. 2009;47(6):647-65. Review.

3) Piroddi M, **Palmese A**, Pilolli F, Amoresano A, Pucci P, Ronco C, Galli F.
   Plasma nitroproteome of kidney disease patients.
   Amino Acids. 2011 Feb;40(2):653-67. Epub 2010 Jul 31.

4) **Palmese A**, De Rosa C, Marino G, Amoresano A.
   Dansyl labeling and bidimensional mass spectrometry to investigate protein carbonylation.
   Rapid Commun Mass Spectrom. 2011 Jan 15;25(1):223-31.

5) Amore A, Amoresano A, Birolo L, Henrissat B, Leo G, **Palmese A**, Faraco V.
   A family GH51 α-L-arabinofuranosidase from Pleurotus ostreatus: identification, recombinant expression and characterization.
   Appl Microbiol Biotechnol. 2011 Nov 13. [Epub ahead of print]

Congress communications

1) Expanding the role of iTRAQ® chemistry: quantitative analysis of post-translational modifications.
   Palmese A, Chiappetta G, Corbo C, Marino G, Amoresano A.
   4th ItPA Congress, Milano, Italy (22-25 June, 2009).

2) Technical advances in proteomics mass spectrometry: identification of post-translational modifications.
   Giangrande C, Amoresano A, Carpentieri A, **Palmese A**, Chiappetta G, Marino G, Pucci P.
   4th ItPA Congress, Milano, Italy (22-25 June, 2009).

3) iTRAQ chemistry to investigate protein carbonylation.
   Palmese A, De Rosa C, Marino G, Amoresano A.
   5th ItPA Congress, Firenze, Italy (9-12 June, 2010).

4) Expanding the role of iTRAQ® chemistry: analysis of post-translational modifications.
Palmese A, Chiappetta G, Corbo C, De Rosa C, Marino G, Amoresano A.
EMBO Practical Course, Uppsala, Sweden (22-24 August, 2010).

Visiting Appointment

September-October 2011

Dott. Joelle Vinh. Spectrométrie de Masse Biologique et Protéomique CNRS, Ecole Supérieure de Physique et de Chimie Industrielles ParisTech, Paris, France.

DOI 10.1002/pmic.200800493 *Proteomics* 2009, *9,* 1524–1537

RESEARCH ARTICLE

Quantitative identification of protein nitration sites

Giovanni Chiappetta[1], Claudia Corbo[1]*, Angelo Palmese[1]*, Gennaro Marino[2] and Angela Amoresano[2]*

[1] Department of Organic Chemistry and Biochemistry, Federico II University of Naples, Naples, Italy
[2] Department of Organic Chemistry and Biochemistry, School of Biotechnological Sciences, Federico II University of Naples, Naples, Italy

Received: June 6, 2008
Revised: July 31, 2008
Accepted: September 26, 2008

Several labelling strategies have been developed targeting specific amino acid residues and/or PTMs. Methods specifically tailored for the qualitative and sometimes quantitative determination of PTMs have emerged. Many research groups have focused their attention towards o-nitrotyrosine residues, developing various methodologies for their identification, while direct quantification has remained elusive. So far the iTRAQ chemistry has been limited to primary amines. Here, we report a new strategy based on the use of iTRAQ reagents coupled to MS analysis for the selective labelling of o-nitrotyrosine residues. This method was proved to lead to the simultaneous localisation and quantification of nitration sites both in model proteins and in biological systems.

Keywords:
ITRAQ / Mass spectrometry / Precursor ion / Protein nitration / Selective labelling

1 Introduction

A number of tagging (or labelling) strategies have been developed to target specific amino acid residues and/or PTMs. These methods enable the enrichment of subfractions *via* affinity cleanup [1–3]. A proposed approach consists in targeting the specific amino acid by an affinity chromatography step (*e.g.* biotin–avidin chromatography). This procedure was used for the isolation of post-translationally modified peptides [4–7]. The presence of phosphorylation, glycosylation, glycation, nitration and specific types of oxidation are examples of PTMs that can be targeted [8, 9].

Stable isotope methods have been introduced into MS-based proteomics to allow relative changes in protein expression to be determined [10]. The principle of these methods is the incorporation of a stable isotope derivative in one of the states to be compared. Stable isotope incorporation shifts the mass of the peptides by a predictable amount. The ratio of analyte between two or more states can then be determined accurately by the measured peak ratio between the heavy and light derivatised samples One of the first approaches based on the isotope stable affinity tagging and MS is the ICAT strategy [11] specifically addressed towards Cys residues. Another elegant approach use cell-culture enrichment with a stable isotope-labelled amino acid, for *in vivo* incorporation of a mass difference to support relative quantitation [12].

A novel methodology for quantitative analysis by MS makes use of iTRAQ (acronym for isobaric tag for relative and absolute quantification) technology, a newly developed method by Applied Biosystems for relative and absolute quantification of proteins. This novel reagent has been applied for the first time by Pappin and coworkers for the multiplexed protein quantitation in yeast [13]. iTRAQ reagents are specifically reactive towards primary amino groups (namely N-terminal of proteins and peptides and epsilon-amino groups of lysine) and marketed in four different forms called iTRAQ 114.1, 115.1, 116.1 and 117.1,

Correspondence: Dr. Angela Amoresano, Dipartimento di Chimica Organica e Biochimica, Università degli Studi Federico II, Complesso Universitario Monte S. Angelo, via Cynthia 4, 80126 Napoli, Italy
E-mail: angamor@unina.it
Fax: +39-081674313

Abbreviations: N-BSA, nitrated BSA; **PIS,** precursor ion scan; **TEAB,** triethyl ammonium carbonate; **TNM,** tetranitromethane

* These authors contributed equally to the work.

depending on the m/z value of the reporter group. The reagents are differentially isotopically labelled such that all derivatised peptides are isobaric (isobaric tag with a mass of 145.1 Da). In fact, 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. The iTRAQ reagents allow the simultaneously multiplexed analysis of four samples. Each of these samples is labelled with one of the iTRAQ reagents; all peptides with the same sequence, but carrying different versions of the tag will be chromatographically indistinguishable and identical in mass, therefore, also identical in single MS mode. Upon fragmentation, the reporter group is detached, creating signals in the 'quite' low mass range region (114–117 Da) of the MS/MS spectrum. The balance group is lost as a neutral, so that the remaining peptide backbone remains unmodified and can generate backbone fragments that are identical in m/z for all samples, resulting in improved S/N ratios because all differentially coded samples contribute to MS/MS spectrum. Thus, low mass reporter ion signals allow quantitation, while peptide fragment ion signals allow protein identification. Very recently, an octu-plex iTRAQ kit has been proposed thus allowing the simultaneous analysis of eight different samples [14]. Moreover, iTRAQ labelling coupled with pulsed Q dissociation technique has been recently proposed for protein quantitation [15].

Among PTMs, protein nitration can occur in cells during oxidative stress and over-production of nitric oxide [16, 17]. Tyrosine nitration is becoming increasingly recognised as a prevalent, functionally significant post-translational protein modification [18]. This modification is, in fact, implicated in the control of fundamental cellular processes including cell cycle, cell adhesion and cell survival, as well as cell proliferation and differentiation.

Addition of NO_2 group to the ortho-position of tyrosine decreases the pK_a of its hydroxyl group of about three units. This bulky substituent can induce changes in protein conformation, resulting in the generation of antigen epitopes, altered enzyme catalytic activity, modulation of metabolic pathways and inhibition of tyrosine phosphorylation by protein kinases [19–21]. For these reasons and for its chemical stability NO_2Tyr is considered the most important biomarker for identification and quantitation of cellular processes, associated to reactive nitrogen species (RNS) occurrence, that lead to PTM of proteins.

Many research groups have focused their attention towards 3-nitrotyrosine residues, developing various methodologies for revelation and quantification of NO_2Tyr residues [22–25].

So far the iTRAQ chemistry has been limited to primary amines even instrumental for the identification of PTMs. As an example Pflieger and coworkers showed a protocol for phosphorylation quantification [26] based on the combination of iTRAQ isobaric labelling of peptide and phosphatase treatment for the identification of phosphorylated protein in complex mixtures.

We are planning to extend the chemistry of the reagent to address quantification of function of other than primary amines by taking advantage of the experience recently made with dansyl chemistry. In fact, our group has already reported a new approach involving dansyl chloride labelling of nitration sites that rely on the enormous potential of MS^n analysis [27]. A similar approach has been reported by Smith's laboratory for the selective enrichment and identification of nitrotyrosine-containing peptides based on derivatisation with N-succynimidyl S-acetylthioacetate and LC-MS/MS analysis of newly generated amino-tyrosines deriving from sodium dithionite-assisted reduction of nitro groups [8]

Here, we report a new strategy based on the use of iTRAQ reagents coupled to MS analysis for the selective labelling of *o*-nitrotyrosine residues. This method was proved to lead to the simultaneous localisation and quantification of nitration sites both in model proteins and in biological systems.

2 Materials and methods

2.1 Chemicals

BSA, tetranitromethane (TNM), ammonium bicarbonate (AMBIC), guanidine, DTT, trypsin, iodoacetamide (IAM) and triethyl ammonium carbonate (TEAB) were purchased from Fluka. Tris, sodium dithionite as well as the MALDI matrix CHCA were purchased from Sigma. iTRAQ reagents were furnished by Applied Biosystems. Methanol, TFA and ACN are HPLC grade type from Carlo Erba, whereas the other solvents are from Baker. Gel filtration columns PD-10 are from Pharmacia, the HPLC ones from Phenomenex, whereas the prepacked columns Sep-pak C-18 are from Waters.

2.2 BSA nitration

An aliquot of 200 μg BSA was dissolved in Tris 200 mM pH 8.0 buffer and was added to a 830 nmol/mL solution of TNM in ACN in a molar ratio 1:10 BSA/TNM. The reaction was carried out under agitation for 30 min at room temperature. The mix of BSA and nitrated BSA (N-BSA) was then desalted by SEC on a SEPHADEX G-25M column; the elution was carried out with Tris 300 mM and the fractions analysed through UV spectrometry at 220 and 280 nm wavelength. Positive fractions were collected and dried.

2.3 Reduction and carbamidomethylation of BSA

BSA and N-BSA mixture were dissolved in denaturation buffer (guanidine 6 M, Tris 0.3 M, EDTA 10 mM, pH 8.0). Reduction was carried out by using a 10:1 DTT/cysteine molar ratio. After incubation at 37°C for 2 h, iodoacetamide was added to perform carbamidomethylation using an excess of alkylating agent of 5:1 respect to the moles of thiolic

groups. The mixture was then incubated in the dark at room temperature for 30 min. The alkylation reaction was stopped by addition of formic acid, in order to achieve an acidic pH. The product was purified by SEC. The elution was performed with TEAB 50 mM pH 8.0.

2.4 BSA digestion

Protein digestion was carried out in TEAB 50 mM pH 8.0 buffer using trypsin at a 50:1 protein/trypsin mass ratio. The sample was incubated at 37°C for 16 h. Then the sample was dried.

2.5 Acetylation of peptide mixture

Peptide mixture was dissolved in 30% TEAB 500 mM pH 8.0, 70% ACN. Acetylation was performed by using acetic acid N-hydroxy succinimide esther in molar ratio 500:1 reagent/peptides. The reaction was carried out for 1 h at room temperature.

2.6 Nitro groups reduction

Reduction of nitro groups to amino groups was performed by adding. $Na_2S_2O_4$ at 100:1 $Na_2S_2O_4/NO_2Tyr$ molar ratio. The reaction was carried out at room temperature for 10 min under stable stir. The product was purified by SEP-PAK RP-LC with C18 column. Peptides were eluted using 80% ACN, 20% formic acid 0.1%. Fractions were collected and lyophilised.

2.7 Bacterial strains, growth conditions and protein extract preparation

Escherichia coli K12 strain was grown in aerobic conditions at 37°C in LB medium. After 16 h, bacteria were harvested by centrifugation and resuspended in Buffer Z (25 mM HEPES pH 7.6, 50 mM KCl, 12.5 mM $MgCl_2$, 1 mM DTT, 20% glycerol, 0.1% triton) containing 1 μM PMSF. Cells were disrupted by sonication. The suspension was centrifuged at $90\,000 \times g$ for 30 min at 4°C. After centrifugation the protein concentration of the extract was determined with Bradford assay.

2.8 iTRAQ selective labelling

Sample was dissolved in TEAB 500 mM pH 8.0 and then was divided in two shares and each of these was differentially labelled with alternatively iTRAQ reagent 114 and iTRAQ reagent 117, following iTRAQ protocol furnished by Applied Biosystem.

2.9 Labelling of bovine milk protein extract

A sample of commercially available bovine milk was reacted with a 10 mM solution of TNM for 30 min under agitation at room temperature. Modified milk proteins were purified by precipitation with the Amersham Clean Up kit and dissolved in denaturant buffer. Reduction was carried out by using a 10:1 DTT/cysteine molar ratio. After incubation at 37°C for 2 h, iodoacetamide was added to perform carbamidomethylation using an excess of alkylating agent of 5:1 respect to the moles of thiol groups. The mixture was then incubated in the dark at room temperature for 30 min. The alkylation reaction was stopped by addition of formic acid, in order to achieve an acidic pH. The protein mixture was purified by SEC on a Sephadex G-25 M column equilibrated and eluted with 50 mM TEAB. Protein fractions were concentrated and then digested with trypsin as already described. The resulting peptide mixture was acetylated and the nitro groups were reduced using a 200 mM solution of $Na_2S_2O_4$ for 30 min. The protein mixture was then reacted with ITRAQ as described above.

2.10 NanoLC-MS

Peptide mixture, obtained as previously described, was analysed by LC-MS/MS analysis 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 reverse-phase pre-column cartridge (Zorbax 300 SB-C18, 5×0.3 mm^2, 5 μm) at 10 μL/min (solvent A 0.1% formic acid, loading time 7 min). Peptides were separated on an Agilent reverse-phase column (Zorbax 300 SB-C18, 150 mm × 75 μm, 3.5 μm), at a flow rate of 0.2 μL/min with a 5–65% 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). Nanospray source was used at 2.3 kV with liquid coupling, with a declustering potential of 20 V, using an uncoated silica tip from New-Objectives (od 150 μm, id 20 μm, TD 10 μm). Spectra acquisition was based on a survey Precursor Ion Scanning for the ion m/z 114 and 117. It was performed over a mass range of m/z 400–1400 with Q1 set to low resolution and Q3 set to unit resolution. Precursors were collided in Q2 with a collision energy ramp of 25–65 V across the mass range. Spectra acquisition was based on a survey precursor ion scan (PIS). A positive ion Enhanced Resolution Scan was performed at 250 amu/s to determine the charge state of the ion. 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 and the cycle repeated. The entire cycle duration, including fill times and processing times was less than 5.3 s. Data were acquired and processed using Analyst software (Applied Biosystems) and processed by using MASCOT in-house software.

3 Results and discussion

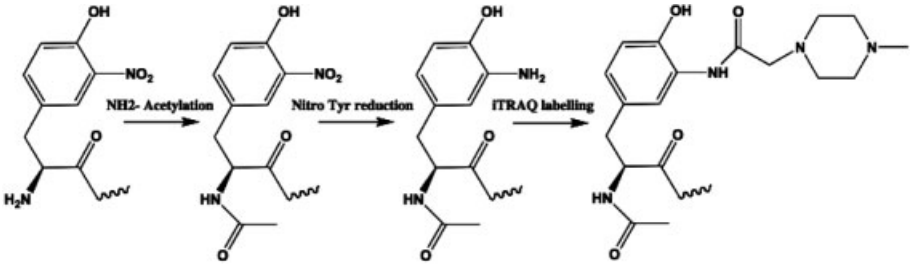
It is well known that some PTMs due to oxidative stress events are linked with the pathogenesis of different diseases, therefore, it is very important to find biological markers able to furnish us information about localisation and physiological effects of oxidative stress. Molecules modified by oxidation, nitration and nitrosilation can act as cellular biomarkers. Identification of the specific targets of protein oxidation was generally accomplished by the detection and quantification on *o*-nitrotyrosine using of 2-DE and Western blot techniques followed by mass spectral identification of protein candidates [28]. However, these procedures generally led to the identification of nitrated proteins but the fine localisation of the nitration sites in complex protein mixture still remain a challenging task. In this respect, our group has suggested a strategy leading to the selective detection of nitrated tyrosine residues by using dansyl chloride labelling coupled to advanced MS experiments [27]. The tryptic digest from the entire protein mixture is directly analysed by MS on a hybrid triple quadrupole/linear ion trap mass spectrometer. Discrimination between nitro- and unmodified peptide is based on two selectivity criteria obtained by combining a PIS and an MS3 analysis. The proposed strategy has resulted to be specific, sensitive, simple and fast. However, because no heavy form of dansyl chloride is commercially available, quantitative analysis of complex protein mixture cannot be addressed. Now taking advantage of the experience made with the above strategy, we used selective labelling of *o*-nitrotyrosine residues exploiting the potential of iTRAQ reagents to perform quantitative analysis of this PTM in proteins. The idea proposed is outlined in the Scheme 1.

According to the so-called 'gel-free procedures', the analysis is carried out at level of peptides following tryptic digest of the whole protein mixture. This approach was first applied to BSA as model protein. 200 μg of BSA were nitrated using TNM as described in Section 2. Then, the resulting BSA and N-BSA mixture was reduced, alkylated and hydrolysed using trypsin. The peptide mixture was then analysed *via* MALDI-MS to verify the extent of nitration.

BSA is a protein of about 66 kDa, showing 20 tyrosine residues along its sequence. However, mass spectral analyses demonstrated that *in vitro* nitration was limited to only few tyrosine residues, namely the ones exposed in the structure of the protein and thus more sensible to nitration. Nitrated peptides were identified by MALDI *mapping* procedure, by comparing the experimental peptides masses with the theoretical values on the basis of amino acid sequence. As already reported, a ΔM of +45 was attributed to the modified peptide, due to introduction of a nitro-group [29]. Figure 1 (panels A and B) showed partial MALDI-MS spectrum of the tryptic BSA and N-BSA mixture. As an example, the signals occurring at m/z 972.5, 1524.8 showed molecular masses 45 Da higher than the theoretical ones. These signals were attributed to the peptides 137–143, 396–408, having the Tyr residue been modified by a nitro group. As a whole, the MALDI-MS analysis showed the occurrence of four signals exhibiting a mass increment of 45 Da. These signals were attributed to the nitropeptides 137–143, 396–408, 322–334, 444–457 (at m/z 972.5, 1524.8, 1612.9, 1770.8, respectively) As further proof, these signals exhibited the characteristic pattern of nitrated peptides, due to photodecomposition of nitrated tyrosine residues by MALDI source [29]. Moreover, mass spectral analysis led to a semiquantitative measure of the degree of the *in vitro* protein nitration. In fact, by comparing the intensity of signals attributed to nitrated peptides with respect to the unmodified one it was possible to estimate that almost 30% of BSA resulted to be nitrated.

In our previous work [27], the selective labelling of *o*-nitrotyrosine residues was essentially based on a reduction step and conversion to *o*-aminotyrosine residues followed by specific dansylation at pH 4.7. A similar approach was attempted here. However, iTRAQ labelling does not permit to exploit the low pK_a value [30] of such amine to obtain a selective labelling, as observed for the dansyl-chloride labelling. Only a small percentage of primary aromatic amine modification by iTRAQ was observed. This behaviour may be explained by the faster rate of hydrolysis of both the ITRAQ *n*-hydroxysuccinimide ester group and the formed aminotyrosine-iTRAQ derivative in mild acidic conditions [31].

Thus, it was necessary to set up a double step procedure. First of all, lysine residues and N-terminus primary amine were protected to facilitate specific labelling of the *o*-nitrotyrosine residues. Blocking can be easily performed under mild conditions by a variety of reagents that are commercially available. In fact, several strategies to introduce NH_2-labels have been described [32, 33]. Generally, it is not possible to reliably control the selectivity of the reaction, because the ε-amino group of lysine is readily and stably modified by reagents that target the N-terminus [34, 35]. Blocking the



Scheme 1. Strategy for selective labelling of *o*-nitrotyrosine residues.

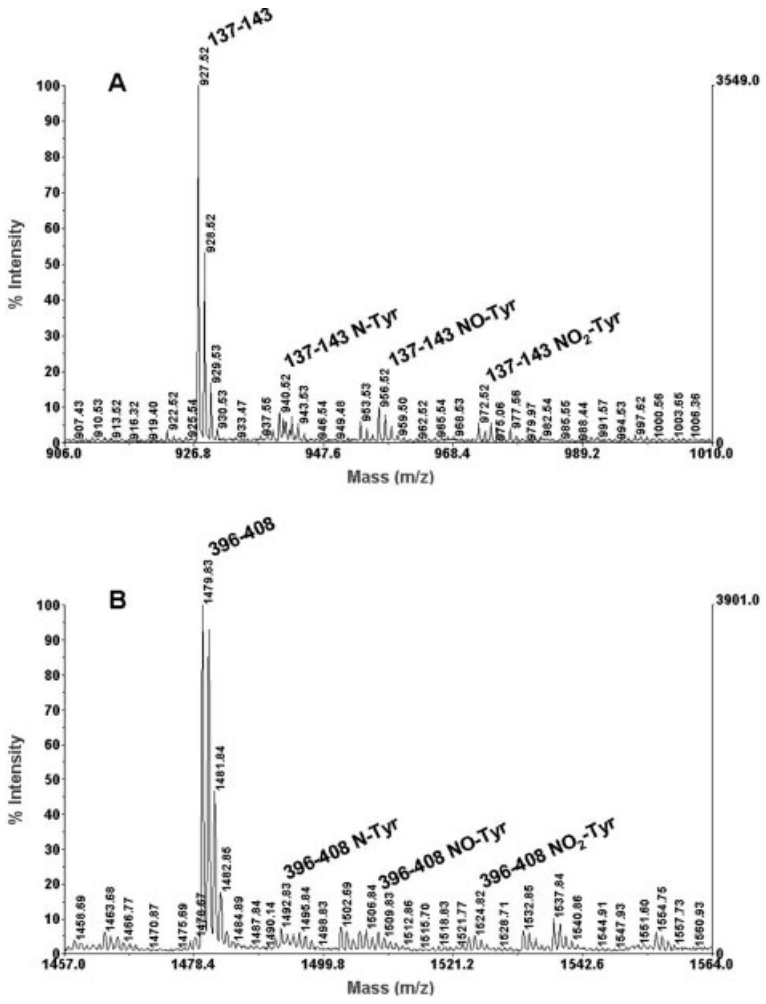


Figure 1. MALDI-MS spectra enlargement of nitrated BSA. (A) Detail of MALDI-MS spectrum relative to nitrated peptide 137–143. (B) Detail of MALDI-MS spectrum relative to nitrated peptide 396–408. MALDI-MS spectrum shows that the degree of protein nitration is almost 30%.

N-terminus and the lysine side chain can be done either prior to or after enzymatic digestion of the sample. By blocking the lysines after digestion, we retained the ability to use trypsin as enzyme for proteolytic cleavage.

Preliminary acetylation experiments were performed by using acetic anhydride as reagent and peptide mixture obtained from trypsin digestion of myoglobin, as model. MALDI-MS analysis performed on the resulting peptide mixture showed that the acetylation reaction was not quantitative. Thus, we decided to use acetic acid *N*-hydroxy succinimide esther as reagent since it has been shown to lead to an efficient lysine acetylation [36]. The yield of the reaction was monitored by MALDI-MS. As an example, Fig. 2 reports the partial MALDI-MS spectrum of the acetylated myoglobin tryptic mixture. The signal at m/z 1649.2 was attributed to the peptide fragment 18–23 carrying an acetyl group, fully modified. No signal related to the unmodified peptide was observed thus indicating that the acetylation was quantitative.

Thus, nitrated peptides were subjected to acetylation using acetic acid *N*-hydroxy succinimide esther as described in Section 2. All successive analysis revealed that every peptide detected had the *N*-term amine and eventually the ε-amino group of lysine protected. Only after the acetylation step it was possible to reduce the nitro-group of *o*-nitrotyrosine residues to an amino group, susceptible to iTRAQ addition. Conversion of nitropeptides into their *amino* derivatives was accomplished by $Na_2S_2O_4$ treatment [30]. The peptide mixture was desalted by sep-pak C18 cartridge and the extent of reduction was monitored by MALDI-MS.

As for quantitative analysis, labelling reaction was carried out on the newly generated amino groups by following the default protocol. iTRAQ technology (trademarked by Applied Biosystems), allows quantitative analysis of protein expression through the use of isobaric tags, enabling the quantitation of four [13] to eight complex protein samples [37] in a single multiplexed analysis. Isobaric tagging methods are emerging as an important tool to study protein expression dynamics.

As an example to illustrate the possibility to use iTRAQ labelling strategy for selective quantitative analysis of *o*-nitrotyrosine residues and at the same time to reduce the costs, we decided to carry out our methodology just on a

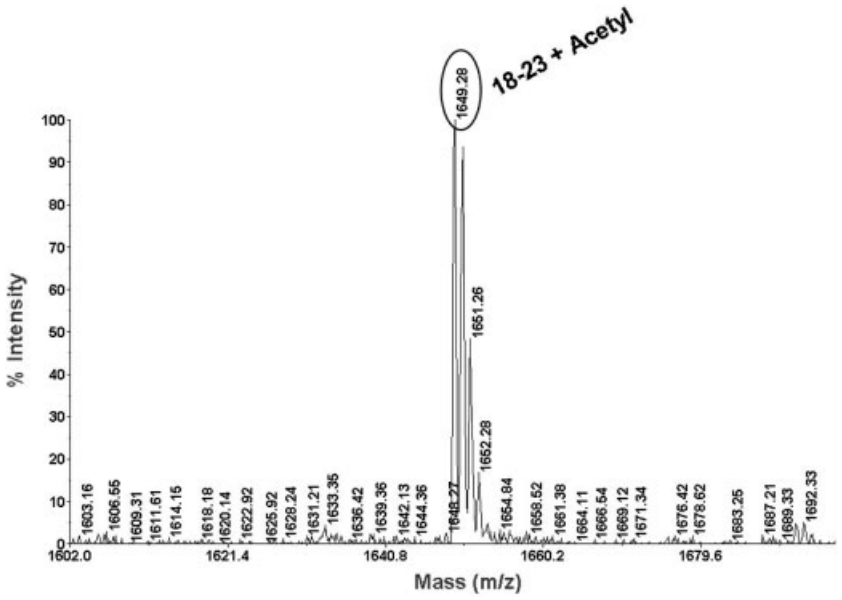


Figure 2. Detail of MALDI-MS spectrum of acetylated myoglobine showing peptide 18–23 with acetylated N-terminal; there are no evidence in the spectrum for non acetylated form of the same peptide.

duplex. Thus, we analysed only two samples and we choose iTRAQ 114 and iTRAQ 117 as labelling reagents. After the labelling, the resulting samples (containing about 30% of N-BSA), N-BSA(114) and *N*-BSA(117), were pooled in molar ratio 1:1 and 1:4, respectively. The two mix (1:1 and 1:4) were then submitted to LC-MS/MS analysis in PIS mode, by using 4000Q-Trap coupled to a 1100 nano HPLC system, thus allowing the simultaneous quantification of differentially labelled peptides in a single chromatographic run. Thus, we setup an experiment by using a PIS for the iTRAQ reporter ions at m/z 117.1.

Moreover, in order to demonstrate that the choice of the fragment to monitor nitrated peptides in PIS mode is irrelevant, because of the isobaric characteristic of the iTRAQ, we decided to analyse the same sample in two separate chromatographic runs, the first in PIS of ion 114.1 and the second in PIS of ion 117.1. As shown in the Fig. 3, the two chromatographic profiles are quite identical and super imposable, differing just in the TIC intensity that reflects the original ratio 1:4 between the two samples. In fact, it should be noted that PIS ion current is recorded from the fragment generated in MS/MS and no memory is retained about the intensity of the parent ion.

To gain information about the sequence of the peptide and the site of nitration, a second stage of mass analysis is performed in which the precursor ion is selected by the first mass analyser and fragmented. Differentially labelled peptides were indistinguishable during a single step of MS analysis due to typical characteristics of iTRAQ reagents. During MS/MS, labelled peptides produce fragmentation spectra that allow the identification of the protein. Moreover, MS/MS produces fragmentation at the sites modified by the iTRAQ reagent yielding reporter ions (m/z 114 and 117) that are visible as separated peaks in the low mass region of the MS/MS spectrum. The ratios of related areas of these peaks were then used to perform a relative quantitative analysis [38-39].

Figure 4 showed the reconstructed ion chromatogram of the 1:1 N-BSA sample analysed by PIS of 117 (A) and a classical full scan MS (B). As clearly shown, the PIS analysis led to reduce the number of undesired signals, improving also the duty cycle of the method and the sensitivity towards nitro-peptides enhancing S/N ratio. However, others signals appear in the chromatogram not related to nitrated peptides. This effect might be generated by the presence of different fragment ions in the region 113-114 or 116-117, many of which generated after acetylation step (Table 1). Thus we used 114 and 117 iTRAQ reagents, in order to have the worst condition of selectivity and demonstrate the feasibility of our methodology.

An accurate manual analysis of MS/MS spectra, revealed that many positive false were selected for the presence of m/z 114 fragment whose intensity would be negligible on quantitative estimation (Fig. 5). To enhance the selectivity of the analysis made by 4000QTrap, we also tried to introduce an MS3 event as done in a previous work [27]. In particular,

Table 1. Unlabelled BSA peptides and their precursor detected in PIS 114 and PIS 117 mode

Interfering ions	Mass	Peptides	Scan
(Ac)b1 Ala	114.1	(Ac)ADE(Ac)K(Ac)KFWGK(Ac)	PIS 114
(Ac)b1 Ala	114.1	(Ac)AEFVEVTK(Ac)	PIS 114
(Ac)Immonium Val	114.1	(Ac)VLASSARQR	PIS 114
(Ac)Immonium Val	114.1	(Ac)VH(Ac)KECCHGDLLECADDR	PIS 114
(Ac)Immonium Thr	116.1	(Ac)TCVADESHACCEK(Ac)	PIS 117
(Ac)Immonium Thr	116.1	(Ac)TVMENFVAFVDK(Ac)	PIS 117

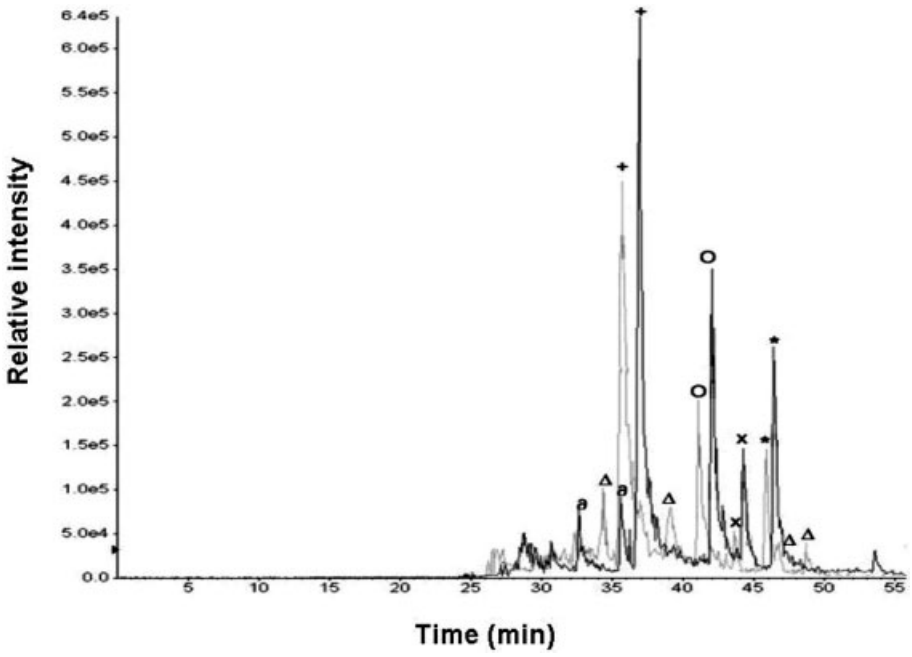


Figure 3. Overlapped chromatograms of nitrated BSA recorded by PIS analysis of ions 114 (grey line) and 117 (bold line), respectively. The chromatogram of PIS 114 is shifted back of 1 min to have a better comparison of signals. The symbols +, ○, X, *, are referred to BSA-nitrated peptides 137–143, 396–408, 444–457, 322–334, respectively. The symbols Δ, ∂ are referred to BSA unspecific peptides detected in PIS of 114 and PIS of 117, respectively.

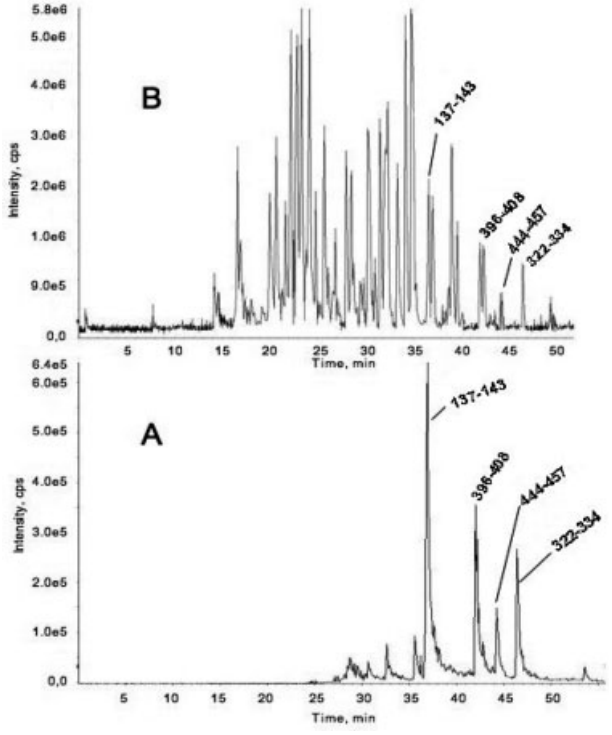


Figure 4. (A) Chromatogram of nitrated BSA recorded in PIS 117. (B) Chromatogram of nitrated BSA recorded in a full scan MS mode.

we tried to fragment in MS/MS the ion at m/z 145 and monitor the loss of CO giving rise to the peculiar reporter ions of iTRAQ in MS3 mode. A Preliminary experiment revealed that MS3 spectra exhibited a very poor intensity caused by the low yield of ion fragment at m/z 145 in MS2 mode. Thus, we found that to enhance the selectivity of LC-MS/MS analysis good results were obtained increasing the signal threshold of PIS using the great difference in signal intensity between the iTRAQ reporter ions and others mentioned previously (Table 1).

However, we inferred that our LC-MS/MS method for 4000QTrap analysis resulted to be a good compromise between selectivity and sensibility that are fundamental parameters in nitro-proteome analysis because of the sub-stoichiometric characteristic of protein nitration. In fact, we want to underline that, without any chromatographic nitropeptide enrichment, just using the mass spectrometer gas phase fractionation features, we had eliminated about 96% of BSA undesired peptides detected with a classical LC-MS/MS approach. Furthermore if eventually, in a borderline case, the detection of a false positive may complicate the analysis, the sample may be re-analysed inverting the diagnostic iTRAQ reporter ion monitored by PIS analysis. Indeed unspecific ions detected in PIS of 114 were not detected in PIS of 117 and *vice versa*.

Identification of nitrated tyrosine residues was carried out by taking advantage of the flexibility of the in house MASCOT software. The software was properly modified by adding in the 'modification file' the value of ΔM 159 for iTRAQ labelled ammino-tyrosine residue. For MASCOT analysis it was necessary to set the software supplying key information, *e.g.* proteolytic enzyme, fixed modifications, variable modifications, sample taxonomy, eventually missed cleavage, *etc.* The data obtained from MS/MS analysis in PIS mode where exploited both to identify nitration

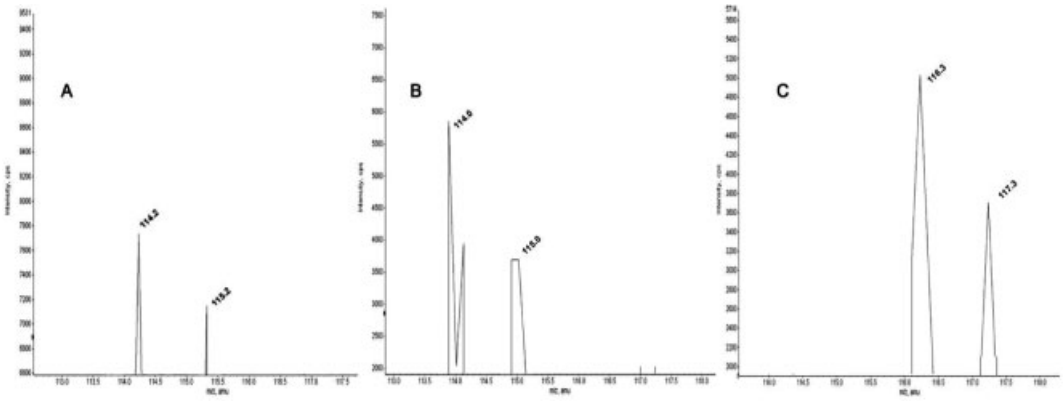


Figure 5. MS/MS spectra enlargement of characteristic iTRAQ *m/z* region of nonspecifically detected BSA peptides. (A) Peptide AEFVEVTK was detected in PIS 114 due the presence of an acetylated b1 ion of alanine residue at *m/z* 114.2. (B) Peptide VLASSARQR was detected in PIS114 mode because of the presence of an acetylated immonium ion of valine at *m/z* 114.0. (C) Peptide TVMENFVAFVDK was detected in PIS 117 mode because of the presence of an acetylated immonium ion of threonine at *m/z* 116.3 or its isotope.

sites by using SwissProt database and to perform quantitative analysis.

Nitration sites Y^{137}, Y^{406}, Y^{457} and Y^{337} were identified by the detection and by MS/MS spectra interpretation from MH_2^{2+} ions at *m/z* 564.8, 841.0, 963.5, 885.3, respectively. As an example, the MS/MS spectrum of modified peptide 137–143 is shown in Fig. 6A. The derivatised product ions were identified from either the corresponding b or y product ion series. By using the Analyst software tool for calculation of peaks area (Applied Biosystems) it was possible to compare the areas signals relative to reporter ions at *m/z* 114 and 117. As an example, for the MS/MS spectrum mentioned above a ratio of 1.1 $\pm$ 0.2 was calculated.

Figure 6B showed the MS/MS spectrum relative to the peptide 444–457 for the mixture 1:4 molar ratios between N-BSA labelled with iTRAQ 114 and N-BSA labelled with iTRAQ 117. In this case nitration site was detectable from y ions and the area ratios for the reporter ions at *m/z* 117 and 114 resulted to be 4.1 $\pm$ 0.4, thus providing an accurate measurement of the relative abundance of each nitrated peptide.

Moreover, the presence of an intense iTRAQ-labelled amino tyrosine immonium ion at 292.1 *m/z* could be detected in both spectra (Fig. 6). This ion might be used as an indication for the occurrence of a nitration site in the peptide.

3.1 Analysis of complex mixtures

In a proof of principle, to investigate the feasibility of applying the method to proteomic analysis, a model system was constituted by adding 5 ng of protein mix containing BSA and N-BSA to 30 mg of an *E. coli* entire soluble protein extract. Aliquots of mix of BSA and N-BSA, as control, and *E. coli* extract spiked with BSA mix were fractionated by SDS-PAGE and submitted to a western blot analysis using anti-nitrotyrosine antibody as indicated in Fig. 7. The analysis showed that the amount of N-BSA was quite impossible to differentiate from the other bacterial proteins. Moreover, in lane D is possible to appreciate that the sample used could be considered similar to a 'dirty' N-BSA *E. coli* protein extract, thus being a good candidate for nitro-proteome study. The total protein extract was then submitted to the procedure described above. The mixture was hydrolysed with trypsin and lysine residues were protected by acetylation. Nitrotyrosines were then reduced with dithionite, and the newly generated amino Tyr-containing peptides were submitted to the iTRAQ labelling protocol, using 114 iTRAQ molecule.

After labelling reaction, LC-MS/MS analysis of the mixture was performed by using 4000Q-Trap, as already described. Figure 8 showed a 3-D graphic, defined as Counter Plot, representing LC-MS/MS parameters, namely elution time, *m/z* and ion intensity expressed in greyscale intensity. This graphical view, obtained by the Analyst software tool, showed the duty cycle and sensitivity improvement of the strategy proposed. A full scan MS analysis of N-BSA spiked with an entire digest of *E. coli* proteins extract is reported Fig. 8A. As clearly shown in figure, this MS analysis is not competitive for such complex peptide mixture. Indeed, due to the large time spent by the spectrometer analysing the most abundant peptides, only three out of four nitrated peptides present in the mixture were detected. The high selectivity of the optimised PIS analysis (Fig. 8B) greatly reduced the number of analysed species, leading to a real increase in sensitivity. Protein Identification of nitrated proteins was carried out using the modified MASCOT software as described above. These peptides corresponded to the same peptides previously detected in the analysis of homogeneous nitrated BSA. Thus, our results underscored the ability of the proposed selective labelling strategy to discriminate nitropeptides from their unmodified counterparts in a complex matrix.

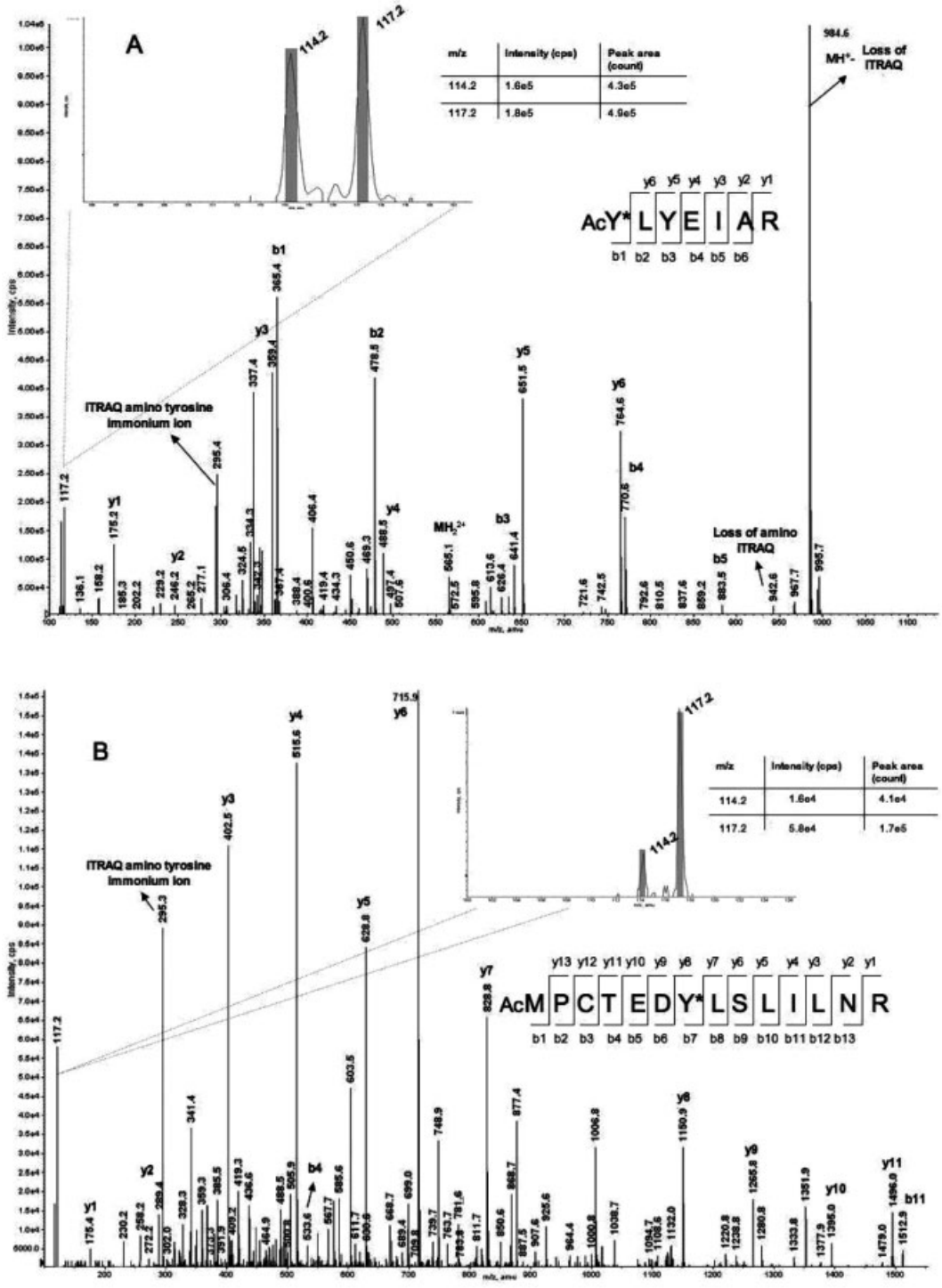


Figure 6. (A) MS/MS spectrum of nitrated BSA peptide 137–143 labelled by iTRAQ reagent in 1:1 ratio. (B) MS/MS spectrum of nitrated BSA peptide 444–457 labelled by iTRAQ reagent in the ratio 1:4. Spectra enlargement showed iTRAQ signals allowing quantitative estimation.

The proposed strategy was finally employed to identify unknown *o*-nitrotyrosine residues in a complex sample as the entire bovine milk. Bovine milk was *in vitro* nitrated with TNM [27] and the extent of nitration was monitored by SDS-PAGE and western blot analysis as described above. Figure 9A (lane A) showed that the high difference of milk protein abundance in the sample reflected the high dynamic range features of the proteomic analysis. In addition the immuno-

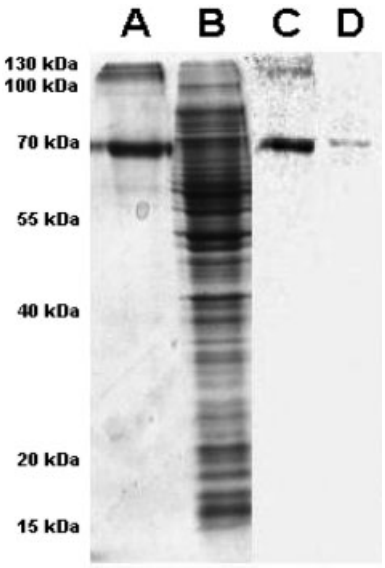


Figure 7. SDS-PAGE and Western blot analyses of *E. coli* protein digest spiked with nitrated BSA (lane B) where N-BSA band was indistinguishable by the others. Lane A: Standard mix of BSA and N-BSA as control. Lanes C and D: Western blot analysis by using antinitrotyrosine antibody of control and *E. coli* samples, respectively.

detection (lane B) showed the low extent of protein nitration thus indicating that this model system could be suitable for proteomic analysis. The entire milk protein extract was then dissolved in denaturant buffer, and cysteine alkylation and nitro-groups reduction were performed 'one pot' as described in Section 2. Two protein samples were then desalted and digested with trypsin. The resulting peptide mixtures were acetylated, differentially labelled with iTRAQ reagents (114 and 117) and mixed in molar ratios of 1:1 and 1:4. The samples were then selectively analysed by nanoLC-MS/MS in PIS mode for the reporter ion at m/z 117.

Identification of nitrated milk proteins was carried out using the modified MASCOT software as described. The results obtained were mediated on triplicate analyses and are summarised in the Table 2. As indicated in the table, besides the high-abundant milk proteins (α-casein and β-lactoglobulin), this procedure was also able to assess the nitration sites occurring in low-abundant proteins, like Vimentin and Integrin, thus, demonstrating the feasibility of this strategy for the identification of protein nitration in proteomics. As for the quantifications, we realised the quantitative estimation of 7/9 nitrated peptides. In fact, the spectra of low abundant proteins let us to unequivocally identify the sequence; however, the quality of signals for the ions 114 and 117 were too low to make an accurate quantitative measure all over the signals. In Fig. 9B the MS/MS spectrum of a nitrated peptide within α-casein is shown. It is interesting to note that the occurrence of an acetylated threonine residue at N-terminal position does not effect the quantitative estimation.

Finally, the proposed strategy was successfully applied to the study of complex mixture of proteins in low abundance, by using plasma samples where real protein nitration changes are observed, thus representing a critical evaluation of this methodology.

Table 2. Nitrated milk proteins detected by the selective iTRAQ labelling strategy coupled with a selective PIS-MS analysis

Observed *m/z*	Score	Sequence	Ratio 114:117	Protein, accession no.
740.2	59	(Ac)VLVDTY*K(Ac)K(Ac)	4.1 ± 0.1 1.2 ± 0.2	β-Lactoglobulin, P02754
796.5	49	(Ac)DMPIQAFLLY*QEPVLGPVR	3.8 ± 0.4 0.9 ± 0.2	β-Casein, P02666
573.5	44	(Ac)TVY*QHQK(Ac)	4.3 ± 0.2 1.2 ± 0.3	α-s1-Casein, P02663
568.8	42	(Ac)Y*LYEIAR	3.9 ± 0.1 0.8 ± 0.3	Albumin bovine, P02769
426.8	36	(Ac)ADLIAY*LK(Ac)K(Ac)	4.1 ± 0.4 1.3 ± 0.2	Cytochrome,*c* P62849
716.3	34	(Ac)STRTVSSSSY*R	4.1 ± 0.4 1.2 ± 0.4	Vimentin, P48616
617.4	30	(Ac)AY*PTPARSK(Ac)	4.2 ± 0.7 1.4 ± 0.3	Glycoprotein hormones α chain, P01217
565.9	29	(Ac)HFHLY*GR	5.2 ± 1.6 1.8 ± 0.9	Integrin α-L, P61625
559.8	29	(Ac)LELY*LPK(Ac)	7.1 ± 3.4 1.8 ± 1.2	Plasma serine protease inhibitor, Q9N212

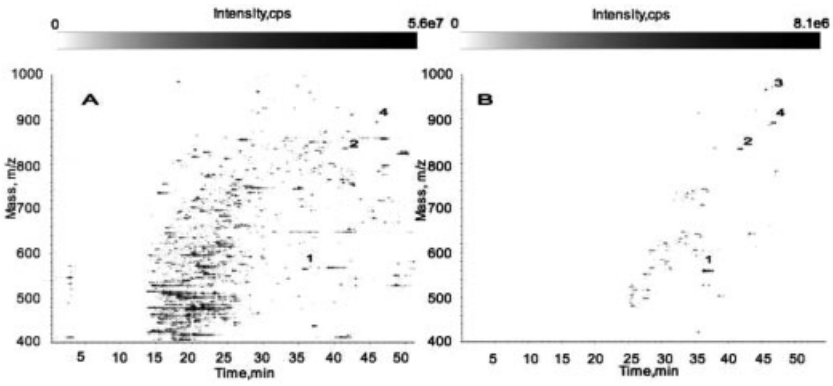


Figure 8. Counter plot representation of a whole *E. coli* protein digest spiked with nitrated BSA analysed, by a classical full scan MS (A) and PIS (B) after iTRAQ selective labelling.

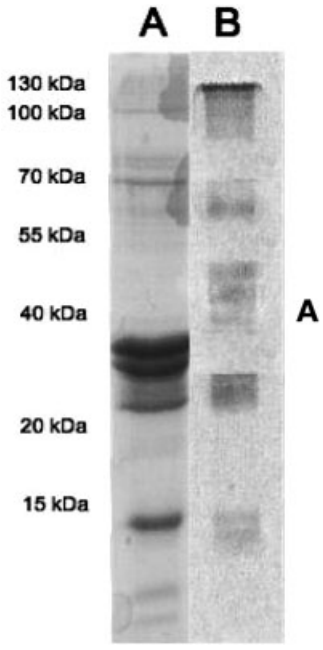
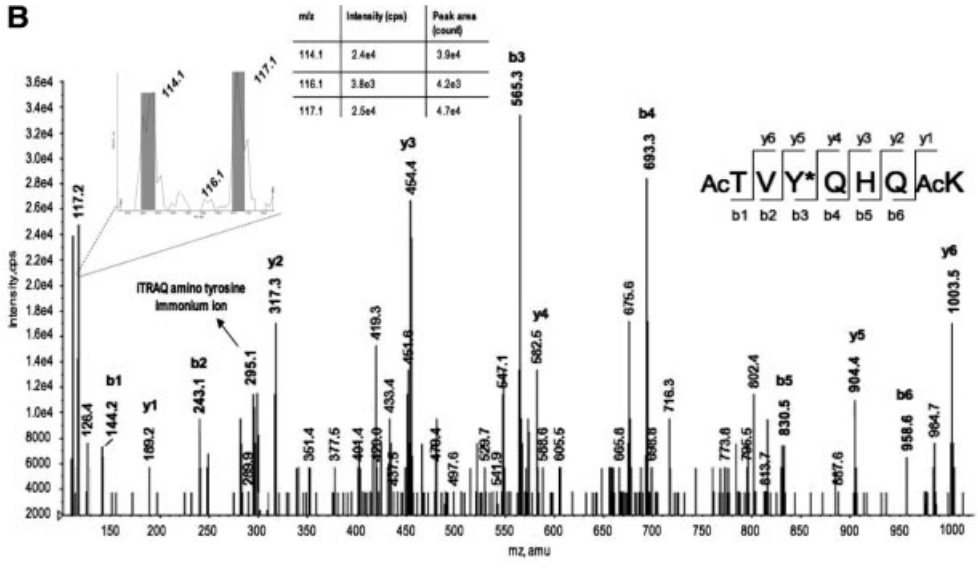


Figure 9. (A) SDS-PAGE (lane A) and Western blot analysis using anti-nitrotyrosine antibody (lane B) of *in vitro* nitrated milk proteins. (B) MS/MS spectrum of α-Casein nitrated peptide TVYQHQK detected in an *in vitro* nitrated milk proteins extract. In the spectrum enlargement of iTRAQ mass region, the characteristic signals of the reporter ions and a signal at *m/z* 116.1 of acetylated threonine immonium ion are visible.



Oxidative and nitrosative stress may cause post-translational changes in plasma proteins that have been described to accumulate in inflammatory and degenerative conditions. Inflammation and oxidative/nitrosative stress (O/NS) are exacerbated in uremia and hemodialysis (HD) patients [40]. Proteomics techniques, also described with the name of 'redoxomics' [41], have been tentatively used in these patients to identify specific PTMs related with O/NS in plasma proteins, and 3-nitro tyrosine is between the most investigated ones.

Thus, the iTRAQ labelling procedure was successfully applied to the qualitative and quantitative analysis of 3NT-containing proteins with different relative abundance and molecular mass, in 100 μL of plasma and dialysis fluids of HD patients (Fig. 10). Our strategy let us to identify the number and position of Tyr residues affected by this post-translational change in individual proteins. Moreover, iTRAQ selective labelling of nitration sites showed that healthy control plasma contained much less nitrated peptides and therefore nitrated proteins, with respect to the uremic ones, and these findings were in agreement with lower levels of biochemical indices of oxidative stress and inflammation (manuscript in preparation).

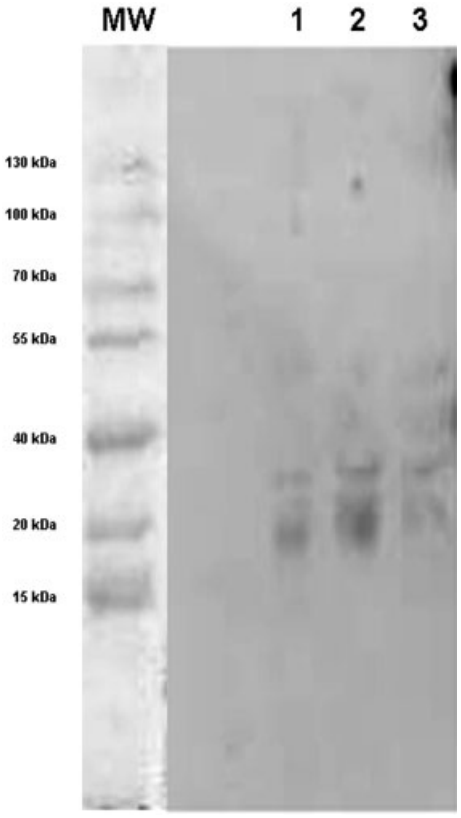


Figure 10. Western blotting analysis of SDS-PAGE fractionation of 3NT-containing proteins in 100 μL of pooled samples of uremic and healthy control plasma (15 ug of proteins/lane). MW, molecular weight markers stained by ponceau; 1, control plasma; 2, uremic plasma, 3, uremic ultrafiltrate.

4 Concluding remarks

The iTRAQ approach has been successfully applied to a variety of prokaryotic and eukaryotic samples including *E. coli*, yeast, human saliva, human fibroblasts and mammary epithelial cells to identify and quantify the proteins in these samples [42–43]. Zhang *et al.* [44] combined the iTRAQ labelling approach with immunoprecipitation to quantify tyrosine-phosphorylated peptides epithelial cells.

However, iTRAQ has been also applied to identify and quantify protein level in shotgun proteomics approach [45–46]. This work represents the first innovative application of iTRAQ labelling strategy to the analysis of PTM. In particular, we addressed the analysis of protein nitration, thus providing a powerful analytical method to identify nitrated protein, localise tyrosine nitration sites and to quantitate the extent of protein nitration in a single experiment.

The newly developed procedure was successfully tested on complex biological systems (manuscript in preparation), to assess its feasibility for proteomic investigation of nitroproteins.

Here, the strategy proposed, taking in account the availability of linear trap to select specific labelled peptides giving rise to diagnostic MS/MS product ions, resulted to be of interest in the selective detection of protein nitration. Moreover, because of its operational simplicity, avoiding long-lasting and time-consuming fractionation procedures, this new strategy seems to be well suited for large-scale quantitative profiling of nitration sites. Because of the cost of iTRAQ kits, we would suggest to integrate our previously proposed method to selectively identify nitration sites on a screening scale and further select the samples to be of interest for quantitative analysis. The strategy reported here was set up with a 4000QTrap mass spectrometer; however, we believe that this approach may be realised also using others mass spectrometer optics, with opportune modifications. For example, Niggeweg *et al.* [47] showed the possibility to perform PIS compatible with chromatographic separations by a QTOF. Moreover, it would be very interesting to evaluate possible improvements, in terms of selectivity and quantitative estimation, by measuring with high resolution and accuracy (<10 ppm) iTRAQ reporter ions. This may be accomplished by a Thermo LIT/Orbitrap performing MS/MS experiments in higher energy CID (HCD) [15]. The optic and the electronics of this instrument leads to perform analysis in an operation mode called 'Parent Ion Mapping' that may be considered as reconstructed PIS. However, our feeling is that, because of the higher duty cycle of a triple quadrupole-based PIS, eventual improvement would have real effectiveness in an offline LC-MALDI-MS/MS analysis more than a classical LC-ESI-MS/MS.

The authors have declared no conflict of interest.

5 References

[1] Lu, Y., Bottari, P., Aebersold, R., Turecek, F., Gelb, M. H., Absolute quantification of specific proteins in complex mixtures using visible isotope-coded affinity tags. *Methods Mol. Biol.* 2007, *359*, 159–176.

[2] Foettinger, A., Leitner, A., Lindner, W., Selective enrichment of tryptophan-containing peptides from protein digests employing a reversible derivatization with malondialdehyde and solid-phase capture on hydrazide beads. *J. Proteome Res.* 2007, *6*, 3827–3834.

[3] Amini, A., Chakraborty, A., Regnier, F. E., Simplification of complex tryptic digests for capillary electrophoresis by affinity selection of histidine-containing peptides with immobilised metal ion affinity chromatography. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 2002, *772*, 35–44.

[4] Temple, A., Yen, T. Y., Gronert, S., Identification of specific protein carbonylation sites in model oxidations of human serum albumin. *J. Am. Soc. Mass Spectrom.* 2006, *17*, 1172–1180.

[5] Camerini, S., Polci, M. L., Restuccia, U., Usuelli, V. *et al.*, A novel approach to identify proteins modified by nitric oxide: The HIS-TAG switch method. *J. Proteome Res.* 2007, *6*, 3224–3231.

[6] Jalili, P. R., Sharma, D., Ball, H. L., Enhancement of ionization efficiency and selective enrichment of phosphorylated peptides from complex protein mixtures using a reversible polyhistidine tag. *J. Am. Soc. Mass Spectrom.* 2007, *18*, 1007–1017.

[7] Mirzaei, H., Regnier, F., Structure specific chromatographic selection in targeted proteomics. *J. Chromatogr. B* 2005, *817*, 23–34.

[8] Zhang, Q., Qian, W. J., Knyushko, T. V., Clauss, T. R. *et al.*, A method for selective enrichment and analysis of nitrotyrosine-containing peptides in complex proteome samples. *J. Proteome Res.* 2007, *6*, 2257–2268.

[9] Zhang, H., Li, X. J., Martin, D. B., Aebersold, R., Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat. Biotechnol.* 2003, *21*, 660–666.

[10] Chen, X., Sun, L., Yu, Y., Xue, Y., Yang, P., Amino acid-coded tagging approaches in quantitative proteomics. *Expert Rev. Proteomics* 2007, *4*, 25–37.

[11] Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F. *et al.*, Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* 1999, *17*, 994–999.

[12] Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B. *et al.*, Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* 2002, *1*, 376–386.

[13] Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B. *et al.*, Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* 2004, *3*, 1154–1169.

[14] Ow, S. Y., Cardona, T., Taton, A., Magnuson, A. *et al.*, Quantitative shotgun proteomics of enriched heterocysts from Nostoc sp. PCC 7120 using 8-plex isobaric peptide tags. *J. Proteome Res.* 2008, *7*, 1615–1628.

[15] Bantscheff, M., Boesche, M., Eberhard, D., Matthieson, T., Robust and sensitive iTRAQ quantification on an LTQ Orbitrap mass spectrometer. *Mol. Cell. Proteomics* 2008, *7*, 1702–1713.

[16] Ischiropoulos, H., Biological tyrosine nitration: A pathophysiological function of nitric oxide and reactive oxygen species. *Arch. Biochem. Biophy.* 1998, *356*, 1–11.

[17] Eiserich, J. P., Patel, R. P., O'Donnell, V. B., Pathophysiology of nitric oxide and related species: Free radical reactions and modification of biomolecules. *Mol. Aspects Med.* 1998, *19*, 221–357.

[18] Shi, W. Q., Cai, H., Xu, D. D., Su, X. Y. *et al.*, Tyrosine phosphorylation/dephosphorylation regulates peroxynitrite-mediated peptide nitration. *Regul. Pept.* 2007, *144*, 1–5.

[19] Lin, H. L., Myshkin, E., Waskell, L., Hollenberg, P. F., Peroxynitrite inactivation of human cytochrome P450s 2B6 and 2E1: Heme modification and site-specific nitrotyrosine formation. *Chem. Res. Toxicol.* 2007, *20*, 1612–1622.

[20] Lopez, C. J., Qayyum, I., Mishra, O. P., Delivoria-Papadopoulos, M., Effect of nitration on protein tyrosine phosphatase and protein phosphatase activity in neuronal cell membranes of newborn piglets. *Neurosci. Lett.* 2005, *386*, 78–81.

[21] Lennon, C. W., Cox, H. D., Hennelly, S. P., Chelmo, S. J., McGuirl, M. A., Probing structural differences in prion protein isoforms by tyrosine nitration. *Biochemistry* 2007, *46*, 4850–4860.

[22] Tedeschi, G., Cappelletti, G., Negri, A., Pagliato, L. *et al.*, Characterization of nitroproteome in neuron-like PC12 cells differentiated with nerve growth factor: Identification of two nitration sites in alpha-tubulin. *Proteomics* 2005, *5*, 2422–2432.

[23] Söderling, A. S., Ryberg, H., Gabrielsson, A., Lärstad, M. *et al.*, A derivatization assay using gaschromatography/negative chemical ionization tandem mass spectrometry to quantify 3-nitrotyrosine in human plasma. *J. Mass Spectrom.* 2003, *38*, 1187–1196.

[24] Krueger, R. C., Use of a novel double-sandwich enzyme-linked immunosorbent assay method for assaying chondroitin sulfate proteoglycans that bear 3-nitrotyrosine core protein modifications, a previously unrecognized proteoglycan modification in hydrocephalus. *Anal. Biochem.* 2004, *325*, 52–61.

[25] Kanski, J., Hong, S. J., Schöneich, C., Proteomic analysis of protein nitration in aging skeletal muscle and identification of nitrotyrosine-containing sequences *in vivo* by nanoelectrospray ionization tandem mass spectrometry. *J. Biol. Chem.* 2005, *280*, 24261–24266.

[26] Pflieger, D., Jünger, M.A., Müller, M., Rinner, O. *et al.*, Quantitative proteomic analysis of protein complexes: concurrent identification of interactors and their state of phosphorylation. *Mol. Cell. Proteomics* 2008, *7*, 326–346.

[27] Amoresano, A., Chiappetta, G., Pucci, P., D'Ischia, M., Marino, G., Bidimensional tandem mass spectrometry for selective identification of nitration sites in proteins. *Anal. Chem.* 2007, *79*, 2109–2117.

[28] Tyther, R., Ahmeda, A., Johns, E., Sheehan, D., Proteomic identification of tyrosine nitration targets in kidney of spontaneously hypertensive rats. *Proteomics* 2007, *7*, 4555–4564.

[29] Pettersson, A. S., Steen, H., Kalume, D. E., Caidahl, K., Roepstorff, P., Investigation of tyrosine nitration in proteins by mass spectrometry. *J. Mass Spectrom.* 2001, *36*, 616–625.

[30] Sokolovsky, M., Riordan, J. F., Vallee, B. L., Conversion of 3-nitrotyrosine to 3-aminotyrosine in peptides and proteins. *Biochem. Biophys. Res. Commun.* 1967, *27*, 20–25.

[31] Jencks, W. J., Enforced general acid-base catalysis of complex reactions and its limitations. *Acc. Chem. Res.* 1976, *9*, 425–432.

[32] Geng, M., Ji, J., Regnier, F. E., Signature-peptide approach to detecting proteins in complex mixtures. *J. Chromatogr. A* 2000, *870*, 295–313.

[33] Brancia, F. L., Oliver, S. G., Gaskell, S. J., Improved matrix-assisted laser desorption/ionization mass spectrometric analysis of tryptic hydrolysates of proteins following guanidination of lysine-containing peptides. *Rapid Commun. Mass Spectrom.* 2000, *14*, 2070–2073.

[34] Zhang, X., Jin, Q. K., Carr, S. A., Annan, R. S., N-Terminal peptide labeling strategy for incorporation of isotopic tags: A method for the determination of site-specific absolute phosphorylation stoichiometry. *Rapid Commun. Mass Spectrom.* 2002, *16*, 2325–2332.

[35] Zappacosta, F., Annan, R. S., N-terminal isotope tagging strategy for quantitative proteomics: Results-driven analysis of protein abundance changes. *Anal. Chem.* 2004, *76*, 6618–6627.

[36] Scholten, A., Visser, N. F., van den Heuvel, R. H., Heck, A. J., Analysis of protein–protein interaction surfaces using a combination of efficient lysine acetylation and nanoLC-MALDI-MS/MS applied to the E9:Im9 bacteriotoxin–immunity protein complex. *J. Am. Soc. Mass Spectrom.* 2006, *17*, 983–94.

[37] Choe, L., D'Ascenzo, M., Relkin, N. R., Pappin, D. *et al.*, 8-plex quantitation of changes in cerebrospinal fluid protein expression in subjects undergoing intravenous immunoglobulin treatment for Alzheimer's disease. *Proteomics* 2007, *20*, 3651–3660.

[38] Skalnikova, H., Rehulka, P., Chmelik, J., Martinkova, J. *et al.*, Relative quantitation of proteins fractionated by the ProteomeLab PF 2D system using isobaric tags for relative and absolute quantitation (iTRAQ). *Anal. Bioanal. Chem.* 2007, *5*, 1639–1645.

[39] Gafken, P. R., Lampe, P. D., Methodologies for characterizing phosphoproteins by mass spectrometry. *Cell Commun. Adhes.* 2006, 5–6, 249–262.

[40] Galli, F., Protein damage and inflammation in uraemia and dialysis patients. *Nephrol. Dial. Transplant.* 2007, *22*, 20–36.

[41] Dalle-Donne, I., Scaloni, A., Giustarini, D., Cavarra, E. *et al.*, Proteins as biomarkers of oxidative/nitrosative stress in diseases: The contribution of redox proteomics. *Mass Spectrom. Rev.* 2005, *24*, 55–99.

[42] Hardt, M., Witkowska, H. E., Webb, S., Thomas, L. R. *et al.*, Assessing the effects of diurnal variation on the composition of human parotid saliva: quantitative analysis of native peptides using iTRAQ reagents. *Anal. Chem.* 2005, *77*, 4947–4954.

[43] Cong, Y. S., Fan, E., Wang, E., Simultaneous proteomic profiling of four different growth states of human fibroblasts, using amine-reactive isobaric tagging reagents and tandem mass spectrometry. *Mech. Ageing Dev.* 2006, *127*, 332–343.

[44] Zhang, Y., Wolf-Yadlin, A., Ross, P. L., Pappin, D. J. *et al.*, Time-resolved mass spectrometry of tyrosine phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules. *Mol. Cell. Proteomics* 2005, *9*, 1240–1250.

[45] Aggarwal, K., Choe, L. H., Lee, K. H., Shotgun proteomics using the iTRAQ isobaric tags. *Brief Funct. Genomic Proteomic* 2006, *5*, 112–120.

[46] Jorrín, J. V., Maldonado, A. M., Castillejo, M. A., Plant proteome analysis: A 2006 update. *Proteomics* 2007, *7*, 2947–2962.

[47] Niggeweg, R., Köcher, T., Gentzel, M., Buscaino, A. *et al.*, A general precursor ion-like scanning mode on quadrupole-TOF instruments compatible with chromatographic separation. *Proteomics* 2006, *6*, 41–53.

This article has been previously published online with preliminary pagination, which has now been corrected.

Clin Chem Lab Med 2009;47(6):xxx-xxx © 2009 by Walter de Gruyter • Berlin • New York. DOI 10.1515/CCLM.2009.154

Review

Technical advances in proteomics mass spectrometry: identification of post-translational modifications

Angela Amoresano, Andrea Carpentieri, Chiara Giangrande, Angelo Palmese, Giovanni Chiappetta, Gennaro Marino and Piero Pucci*

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

Abstract

The importance of post-translational modifications (PTMs) of proteins has become evident in the proteomic era as it plays a critical role in modulating cellular function, and can vary in response to different stimuli thereby tuning cellular mechanisms. Assessment of PTMs on a proteomic scale is a challenging task since they are substoichiometric, transient and reversible. Moreover, the amount of post-translationally modified proteins is generally very small when compared to their unmodified counterparts. Existing methodologies for identification of PTMs essentially relies on enrichment procedure to selectively increase the amount of modified peptides. These procedures need to be integrated with sophisticated mass spectrometric methods to enable the identifications of PTMs. Although the strategies developed so far are not optimal, a number of examples will be given where the combination of innovative separation methods along with advanced mass spectrometric analyses provide positive results. These experiences are leading the way for the next generation of proteomic approaches for identification of a wide range of PTMs.
Clin Chem Lab Med 2009;47

Keywords: glycosylation; nitration; phosphorylation; post-translational modifications (PTMs); proteomics; tandem mass spectrometry (MS/MS).

Introduction

Over the past few years following the conclusion of the Human Genome Project, the challenge has again shifted from the gene to protein, giving rise to what has been termed the ''proteomic era''. The aim of this new era is to identify and characterize proteins located within a given organelle, cell or organism, as well as to unravel the in vivo protein pathways (1–3).

*Corresponding author: Prof. P. Pucci, Dipartimento di Chimica Organica e Biochimica, Università degli Studi Federico II, Complesso Universitario Monte S. Angelo, via Cynthia 4, 80126 Napoli, Italy
Phone: +39081674318, Fax: +39081674313,
E-mail: pucci@unina.it
Received October 11, 2008; accepted March 23, 2009

These new goals are not easily achieved because the challenge increases by several orders of magnitude when moving from genome to proteome research. The static nature of the genome cannot be compared to the dynamic properties of the proteome. Expression profiles of proteins change several times during the cell cycle and are significantly influenced by a number of intra- and extra-cellular stimuli including temperature, stress, apoptotic signals, and others (4). In addition, alternative splicing and post-translational modifications (PTMs) has resulted in re-thinking of the paradigm ''one gene-one protein'' that no longer reflects the true nature of the cellular proteome.

Modern proteomics research is focused on two major areas, expression proteomics and functional proteomics. Expression, or differential, proteomics aims to investigate changes that occur in the expression of protein profiles following specific stimuli and/or pathophysiological conditions. The goal here is to assess up regulation and down regulation of protein concentrations in abnormal cells compared with normal cells. Thus, expression proteomic studies are strictly concerned with transcriptomics experiments. However, mRNA concentrations do not necessarily correlate with the amount of corresponding protein due to translation regulation mechanisms. Thus, expression proteomics complements microarray investigations. This comparative approach is usually employed in the identification of proteins that are up regulated or down regulated in specific disease processes and therefore, may be used as diagnostic markers or therapeutic targets (5). It is not enough to map and identify all the proteins in cells; they must also be categorized with respect to their function and interactions with other proteins or genes. Therefore, the proteome must be studied in vivo, and the dynamic properties of the proteome used to obtain this information. This is the target of functional proteomics, an emerging area of proteomic study to investigate protein function in vivo through the isolation of functional complexes within the cell and the identification of their protein components. It is well understood that many cellular processes are governed by the rapid and transient association of protein complexes that fulfill specific biologic functions. The association of an unknown protein with others belonging to a specific complex would be strongly suggestive of its biological function.

In this scenario, the importance of PTMs is notably increased. Protein modification can affect biological function, thereby playing a critical role in cellular function. Also, they can vary in response to environ-

mental stimuli allowing cellular mechanisms to be finely tuned. Deregulation might be involved in the development of disease.

Several mass spectrometry (MS)-based strategies have previously been developed to identify and characterize PTMs. With these techniques, a single, largely homogeneous protein was digested, with the resulting peptide mixture analyzed using MS procedures. The site and nature of the PTM could be easily inferred from the anomalous mass value detected for the modified peptide(s). The analytical problems faced today by proteomic laboratories are far more complicated. First, the sample very often consists of a complex mixture of proteins, some of which are modified. Second, the PTMs involved in the regulation of important cellular processes, such as phosphorylation, methylation, acetylation, nitration, etc., are substoichiometric, affect only a minor portion of the protein, are transient and are reversible. Therefore, the amount of modified peptides within a proteomic digest will generally be very small. Also, they will always occur in the presence of a much larger concentration of the corresponding unmodified fragments. Thus, an understanding of cellular mechanisms and processes requires a comprehensive assessment of PTMs and the changes that occur following signaling events (6). New strategies will have to be employed based on the integration of specific enrichment procedures with sophisticated mass spectrometric experiments.

This paper will review current existing methodologies to identify PTMs in proteomic experiments, with a focus on some specific examples.

Mass spectrometry

MS had been the method of choice to elucidate PTMs in protein chemistry. Today, MS plays a key role in proteomics studies since it is the only method with the sensitivity, accuracy, reproducibility and speed required for identifying unknown proteins and characterizing transient PTMs (7–9).

MS is based on the production of ions in a gaseous phase that allow for the determination of the molecular mass of analytes. A wide variety of mass spectrometers are now available; all sharing the ability to assign mass-to-charge values to ions, even though, the principles of operation and the types of experiments that can be done on these instruments differ greatly. The sample has to be introduced into the ion source of the instrument where molecules are ionized. The ions are extracted into the analyzer region where they are separated according to their mass-to-charge ratio (m/z) and then detected.

Mass spectrometers can be classified according to its ionization system; an essential component that defines the measurable mass range, sensitivity and resolution. Many ionization methods are available. However, the most widespread ion sources in biochemical analyses are electrospray ionization (ESI) and matrix assisted laser desorption ionization (MALDI) (10, 11). Over the last few years, ESI (12, 13) has had a tremendous impact on the use of MS in biological research because it is well suited to the analysis of polar molecules. During standard ESI, the sample is dissolved in a polar, volatile solvent and pumped through a narrow, stainless steel capillary (12–14). High voltage, 3–5 kV, is applied to the tip of the capillary within the ion source. Sample emerging from the capillary tip is dispersed into a spray of highly charged droplets. This process is aided by a nebulizing gas flowing around the outside of the capillary. The charged droplets decrease in size by solvent evaporation which is assisted by a flow of warm nitrogen gas. This is also known as the drying gas and it passes across the front of the ionization source. Eventually, charged sample ions that are free of solvent are released from the droplets and pass into the analyzer of the mass spectrometer. One peculiar aspect of this technique is that it gives rise to multiple, charged molecular-related ions such as $(M+nH)^{n+}$ in positive ionization mode and $(M\text{-}nH)^{n-}$ in negative ionization mode.

MALDI-MS is based on bombardment of sample molecules with a laser to induce sample ionization. The sample is pre-mixed on a stainless steel plate with a highly absorbing matrix such as a weak aromatic acid. Once excited, the matrix is able to transfer energy and protons to the sample. After drying, the matrix molecules crystallize and solid crystals consisting of sample and matrix are formed. A laser, usually a pulsed nitrogen laser at 337 nm, is fired into the sample. The absorbance of the laser energy by the matrix molecules results in a desorption event. This energy is used to transfer a proton from the matrix to the analyte resulting in the formation of MH^+ ions.

Mass analyzers may incorporate quadrupoles (Q), ion traps (IT), time-of-flight (TOF), or some combination of these in ''hybrid instruments'' in order to obtain good resolution and sensitivity. The ESI source is usually employed with quadrupole and ion trap analyzers (15, 16). The advantages of the ion-trap mass spectrometer includes compact size, and the ability to trap and accumulate ions in order to increase the signal-to-noise ratio. It is also possible to use this type of mass spectrometer in fragmentation experiments (see below). Conventional ion trap mass spectrometers operate with a three-dimensional (3D) quadrupole field. This confers very high efficiency with respect to the time required to fill the ion trap and generate a complete mass spectrum. However, there are some problems with respect to trapping efficiency, primarily due to the small volume. These problems have been overcome by the introduction of linear ion traps (LIT) (17–19), characterized by a larger ion storage capacity and greater trapping efficiency (see below).

TOF analyzers are typically used in combination with MALDI sources. Following MALDI ionization, ions pass into a TOF tube for mass analysis. The TOF tube is under high vacuum (10^{-6}–10^{-8} mbar) and consists of a field-free drift region. All ions entering the TOF tube have a fixed kinetic energy that is propor-

tional to the applied voltage and the charge. The higher the mass of the ion, the lower its velocity and the longer it takes before the ion reaches the detector. Ions of different mass can be separated based on their different velocities in the TOF-tube. Their m/z values are determined by measuring the time each ion takes to travel through the field free region.

MALDI and electrospray sources are known as soft ionization methods because the sample is ionized by the addition or removal of a proton, with very little energy remaining to cause fragmentation of the sample ions. Tandem mass spectrometry (MS/MS) must then be used to produce structural information about the compound by fragmenting specific sample ions inside the mass spectrometer, and then identifying the fragment ions. MS/MS uses two stages of mass analysis; the first to pre-select an ion and the second to analyze fragments. These are separated using a collision cell into which an inert gas, such as argon or xenon, is introduced to produce fragmentation. MS/MS also enables specific compounds to be detected in complex mixtures based on their characteristic fragmentation patterns.

Peptides fragment in a well-described manner (20). There are three different types of bonds that can fragment along the amino acid backbone: NH-CH, CH-CO, and CO-NH bonds. Each bond break gives rise to two species, one neutral and the other one charged, with only the charged species monitored by the mass spectrometer. Thus, there are six possible fragment ions for each amino acid residue as shown in Figure 1. The a, b, and c ions carry the charge on the N-terminal fragment, and the x, y and z ions retain the charge on the C-terminal fragment. The most common cleavage sites are the CO-NH bonds which give rise to the b and/or the y series of ions.

Protein identification by mass spectrometry

The key step in any proteomic study consists in the identification of proteins that have been separated either by gel electrophoresis and digested, or cleaved directly using enzymatic procedures in the so-called ''bottom-up'' approach (21). Protein identification from 2D gel electrophoresis is usually obtained by MALDI-MS analyses (22, 23). Following colloidal Coomassie staining, protein components are excised from the gel, reduced and alkylated with iodoacetamide to irreversibly block the cysteine residues and digested with trypsin. The resulting peptide mixture is extracted from the gel and analyzed directly using MALDI-MS (24, 25).

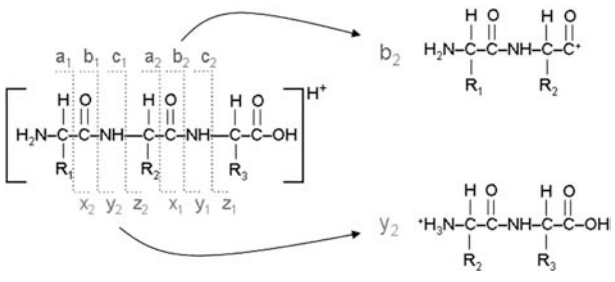


Figure 1 MS/MS peptide fragmentation pattern.
Fragment ions are labeled with a, b, and c ions carrying the charge on the N-terminal fragment. The x, y and z ions having the charge retained on the C-terminal fragment.

Identification of the different proteins is performed using the peptide mass fingerprinting procedure (26). This procedure uses the mass values, together with other information such as the protease used for the hydrolysis. This information is input into the Mascot mass search software or analogous programs available on the internet (27). The experimental mass values are compared to those obtained from a theoretical digestion of all proteins present in the database, allowing for identification of the protein(s).

Alternatively, if the fingerprinting procedure is unable to identify the proteins, combined liquid chromatography-tandem mass spectrometry (LC-MS/MS) methods can be used. Peptide mixtures are fractionated by capillary HPLC. The fractions that elute from the column are applied directly into the electrospray MS and mass values determined. Peptide ions will be simultaneously isolated and fragmented within the mass spectrometer, producing daughter ion spectra from which information on sequence of individual peptides can be obtained. This information, together with the peptide mass values, is then used to search protein databases enabling identification of the protein components. It has been shown that sequence information from just two peptides is sufficient to identify a protein by searching protein and expressed sequence tag databases.

Recently, ''top-down'' proteomics has become popular in the assessment of PTMs. In this approach, the multicharged ions of intact proteins are used as precursor ions and fragmented to provide sequence information and protein identification. Technically demanding, sophisticated instrument setup constitutes the main challenge of top-down proteomics (28). However, recent developments in instrumentation have been introduced, such as specifically designed Fourier transform ion cyclotrone resonance (FTICR) mass spectrometers, making the top-down approach suitable even for PTM studies (29).

Post-translational modifications (PTMs)

Protein PTMs are thought to be one of the major determinants for the complexity of life. These modifications increase the molecular heterogeneity of gene products, adding a further dimension to the proteome as compared to the genome. Greater than 5% of the genes in the human genome encode for enzymes that perform such modifications. These include hundreds of protein kinases and phosphatases, ubiquitinyl ligases, acetylases and deacetylases, methyl transferases and glycosyl transferases (30). The repertoire of PTMs is very complex considering that no <15 out of the 20 common amino acids may be modified. These modifications range from addition of a single group to the linkage of a complex oligosaccharide moiety or glycosyl phospohatidyl

inositol (GPI) anchor. More than 250 PTMs have been described to date (31). Genomic studies have not been able to elucidate complex mechanisms such as development of disease, aging, differentiation, development, and others that are finely regulated by dynamic changes in protein function mediated through PTMs.

The pattern of protein PTMs constitute a molecular code that dictates protein conformation and folding, cellular localization, cellular interactions and cell activities, tissue and environmental conditions. Understanding this code constitutes a major challenge of proteomics. It should be emphasized that detection of PTMs, while important, does not distinguish the number and type of PTMs that may be present on any one intact protein. This disconnection from the parent protein modifications may lead to the loss of information on the synergies between PTMs. New methods and strategies have been developed, and many more are needed to fully characterize PTMs in a specific proteome and to describe their changes.

Glycosylation

Glycosylation is a co- and post-translational modification where oligosaccharide chains are covalently linked to proteins. Glycoproteins are ubiquitous, being found in all living organisms from eubacteria to eukaryotes (32). They fulfill their biological roles as a complex mixture of glycosylated variants called glycoforms. This microheterogeneity is primarily due to the fact that oligosaccharide biosynthesis is not template driven. Assembly of the glycan moieties is governed by the relative amount of different glycosidic enzymes (33), the cell type, and physiological status. In addition, the reactions involved in glycosylation and deglycosylation are frequently incomplete, resulting in an increase in the heterogeneity of glycoproteins.

A new discipline, glycobiology, emerged in the 1980s, with a focus on the structural and functional characterization of glycoproteins and glycoconjugates (34). Recently, glycobiology has evolved into glycomics, the investigation of the complete set of glycoconjugates and carbohydrates that occur in an organism. Glycomics show a higher level of complexity than genomics, or even proteomics, primarily due to the microheterogeneity of glycoforms described above. Furthermore, glycans have a larger number of building blocks compared with nucleic acids and proteins. These can be assembled in complex branched structures in which monomeric units can adopt various conformations and binding positions.

Biosynthesis of glycopeptide linkages

The enzymes responsible for glycosylation are membrane-bound glycosyltransferases and glycosidases that are active in the endoplasmic reticulum (ER) and Golgi compartments. Their activity can generate two large classes of glycoforms; N-linked and O-linked glycans. O-linked oligosaccharides have the glycosidic component linked to the hydroxyl group of serine or threonine residues. N-glycosylation is characterized by the formation of an N-glycosidic linkage on the amide of asparagine residues. This glycosylation requires a strict consensus sequence *Asn-X-Ser/Thr* (35). Although this sequence frequently occurs, it is not always a site that is gylcosylated. This is most probably due to conformational factors. N-glycosylation begins in the ER (36) where a biosynthetic precursor, a dolichol-P-P-linked oligosaccharide ($Glu_3Man_9GlcNAc_2$), is transferred to Asn residues in nascent polypeptide chains. This glycan is a well-conserved structure in evolution; the same compound is seen in plant, fungal and mammalian cells.

Further processing of the oligosaccharide may occur in the Golgi complex. The structural diversity found in Asn-linked oligosaccharides from different organisms, cell types or tissues, originates from processing reactions occurring in the Golgi apparatus (37, 38).

O-glycosylation also occurs in the Golgi and consists of the step-wise addition of sugar residues directly onto the polypeptide chain. This process starts from an N-acetylgalactosamine residue linked to the hydroxyl group of Ser or Thr residues (39). A specific glycosyltransferase then adds the galactose residue to the GalNAc. The last steps in biosynthesis of typical O-glycans are the additions of two N-acetylneuramic acid (sialic acid) residues in the trans-Golgi reticulum. O-oligosaccharides can vary in size from a single GalNAc residue, referred to as the Tn antigen, to oligosaccharides comparable in size to complex N-oligosaccharides.

The role of glycosylation

The oligosaccharide moiety confers specific physicochemical properties to glycoproteins, and fulfills several biological functions. Several studies dealing with the importance of N-glycosylation in the folding process have been reported. Glycans constitute bulky hydrophilic moieties that reduce aggregation and promote solubility. This feature helps the protein follow a specific folding pathway. However, the principal role in the folding process is played by deglycosylation and reglycosylation in the ER, which occurs repeatedly during the biosynthesis of glycoproteins. This system is recognized as a central component of a quality control and chaperone pathway in the ER lumen (37, 40–42).

Some oligosaccharides play an important role in the coating of many glycoproteins, protecting them from proteases and antibodies and affecting their turnover and half-life (43, 44). Oligosaccharides can act as specific receptors for pathogens, being responsible for interactions with vertebrate cells through lectins. Lectins are specialized glycoproteins that can recognize vertebrate glycan structures with high specificity (45). The variable nature of surface glycans may be a way to escape pathogens' recognition systems.

Recognition by ligands or receptors is a constant characteristic of glycoproteins. Most are cell-surface receptors, whose glycosylation affects binding and activity. Cells are covered with a high-density coating of oligosaccharides, indicating that oligosaccharides are responsible for cell-cell and cell-matrix interactions and adhesion (43).

Glycoproteome characterization

Understanding the structure, enzymology and cellular biology of glycosylation is the main goal of glycomics with the ultimate goal being the comparison of organisms in normal and pathological conditions (44). However, difficulties exist in the analysis of complex mixture of glycoconjugates, primarily due to their microheterogeneity as stated above. These difficulties largely prevent the simple and direct approach of analyzing the entire set of glycans following their release from glycoproteins or glycoconjugates. Several protocols (46–49) have been developed to release the oligosaccharides from both O-linked and N-linked glycoproteins. However, the ensemble of all these glycans constitute a far too complex mixture to be analyzed directly by any mass spectrometric method. These analyses can provide information only on the most abundant glycoforms, leaving the vast majority of minor glycans undetected.

Various methods have been developed to enrich glycoproteins in a complex mixture. One approach relys on capturing N-linked glycoproteins using hydrazine chemistry. The oligosaccharides of the glycoproteins are oxidized to form the corresponding di-aldehydes that can be covalently linked to hydrazide beads. The captured glycoproteins are then digested and the unmodified peptides washed out. Captured glycopeptides can be deglycosylated with PNGase F and quantified by isotopic labeling (50).

Chemical derivatization presents several side reactions. Thus, lectin affinity chromatography has recently been developed as an analytical tool in glycoproteomics. Several lectin-derivatized media are commercially available for enrichment of glycoproteins and glycopeptides. A single lectin recognizes specific glycoforms, whereas the use of a set of lectins allows the complete set of glycoproteins to be captured in a single step. Multilectin affinity columns have been developed combining different lectins: Hancock and co-workers combined concanavalin A (ConA), wheat germ agglutinin (WGA) and Jacalin for the analysis of the serum glycoproteins (51). Comparative serum glycoproteomics has been applied to the study of pathological samples using a multilectin affinity approach to identify oligosaccharide markers in such diseases (33, 52–54).

Alternative methods rely on the chemical and/or enzymatic hydrolysis of glycoproteins, followed by affinity capture of the entire ensemble of glycopeptides with multiple lectins. Selective isolation of glycopeptides allows larger recovery of glycoforms since the steric hindrance of glycoproteins severely affect the interaction with lectins.

The structural characterization of glycans is still a difficult task due to the presence, in most cases, of several different structures at the same glycosylation site (46). Unfortunately, although MS cannot differentiate between the different hexoses and hexoseamines, it has become a valuable tool in their characterization (55).

The determination of the entire structure of oligosaccharide chains requires that the intact sugar moieties be released from the glycoproteins/glycopeptides pool using either PNGase F or chemical methods (56–61). Glycan profiling is then achieved by MALDI-MS analysis of the intact glycan mixture. The MALDI-MS spectra show the presence of intense quasi-molecular ions originating from the whole oligosaccharide moieties; assignment of the different glycoforms is then performed on the basis of the molecular weight of the oligosaccharides and their known biosynthetic pattern (46).

This analysis is not sufficient to provide structural information about the sequence and linkages of glycans. Combination of MALDI-profiling with post-source decay (PSD) or collision-induced dissociation (CID) analyses is generally required (62). Glycan fragmentation spectra present some main cleavage sites. A systematic nomenclature for oligosaccharide fragment ions was established by Domon and Costello (63), following that already in use for peptides. However, the structure of carbohydrates necessitates the use of a more complex annotation due to possible breakages occurring within the sugar ring. The primary cleavage sites, termed glycosidic cleavages, involves the breaking of a single bond between adjacent sugars (Figure 2). They provide information on sequence and branching, but little information on linkage. Other cleavage sites, termed cross-ring cleavages involving cleavage of two bonds are useful in determination of the linkage. It is worth noting that ions retaining the charge on the reducing terminus are labeled X (cross-ring), Y ($C_1 \rightarrow$O-glycosidic) and Z (O$\rightarrow C_x$ glycosidic), while if the charge is retained at the non-reducing end, the produced ions are called A, B and C. The B and Y ions are usually the most abundant species.

Cross-ring cleavage necessitates additional superscript prefixes, denoting the atoms of the two cleaved

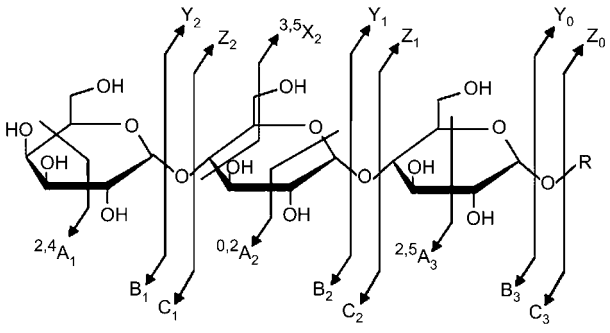


Figure 2 Carbohydrate fragmentation pattern.
Ions retaining the charge at the reducing terminus are named X, Y and Z. Complementary ions are labeled A, B and C.

bonds. The situation becomes more complicated in cases involving branched oligosaccharides. The nomenclature is based on the carbohydrate component of glycopeptides and glycoproteins, which consist of a ''core'' unit and the branches, usually called ''antenna''. Each antenna is identified by a Greek letter with the largest being given the symbol ''α'' and the other symbols assigned based on decreasing molecular weight.

CID provides focused ions and nanospray technology provides good sensitivity. It is possible to perform MS/MS or MS^3 experiments in order to obtain extensive fragmentation of oligosaccharides or glycopeptides. This analysis helps elucidate the different glycosidic structures present in glycoproteins including the branching pattern, the sequence of antenna and the possible occurrence of modifying groups, such as sulfate, phosphate, acetyl groups, and others (64).

Glycoproteome in tumor and inflammation: a case study

Recently, our group introduced a simple strategy that provides preliminary insight on the comparison of glycoproteomes in sera from healthy and diseased individuals. This technique uses a single affinity chromatography step coupled with MS (Figure 3). The N-linked glycopeptides resulting from digestion of sera samples with trypsin were selectively enriched in order to enhance identification of N-glycosylation sites using LC-MS/MS, and to obtain glycoform profiles using MALDI-MS analysis. Human sera from patients with either myocardial damage or hepatic carcinoma were evaluated since they represent examples of inflammatory and malignant disease.

Myocarditis is a group of diseases caused by infection, toxins, or autoimmune mechanisms characterized by inflammation of the heart, leading to necrosis of myocytes (65, 66). Glycan moieties might be responsible for the increase in leukocytes in damaged tissue. Hepatic carcinoma is a primary malignancy of the liver. There are several hypotheses about the role that glycans play in the development of cancer. C glycan structures are known to contribute to oncogenic signaling, transformation and metastasis (67).

Aliquots of serum from the two different patient groups, and healthy donors, were pooled and digested with trypsin. The three peptide mixtures were subjected to ConA affinity purification. The glycopeptides were then deglycosylated using PNGase F treatment and analyzed by MS. Similar analysis was performed on the unbound ConA fractions which contained primarily non-glycosylated peptides and highly branched glycopeptides. The ConA affinity step specifically retained the high mannose, hybrid and biantenna structures, separating them from the rest of the glycans.

The oligosaccharides retained by ConA were eluted and the glycans analyzed with MALDI-MS. Each signal in the MALDI spectra was tentatively attributed to a glycan arrangement on the basis of the mass value and the known N-glycan biosynthetic pathway. Since desialylation might have occurred in both positive and negative ion mode, MALDI analyses were also performed using 2,4,6 trihydroxyacetophenone as matrix. This has been reported to minimize loss of sialic acids (68).

More than 50 arrangements were detected, belonging primarily to high mannose and biantenna classes, indicating a high degree of glycan heterogeneity. While some qualitative and quantitative differences

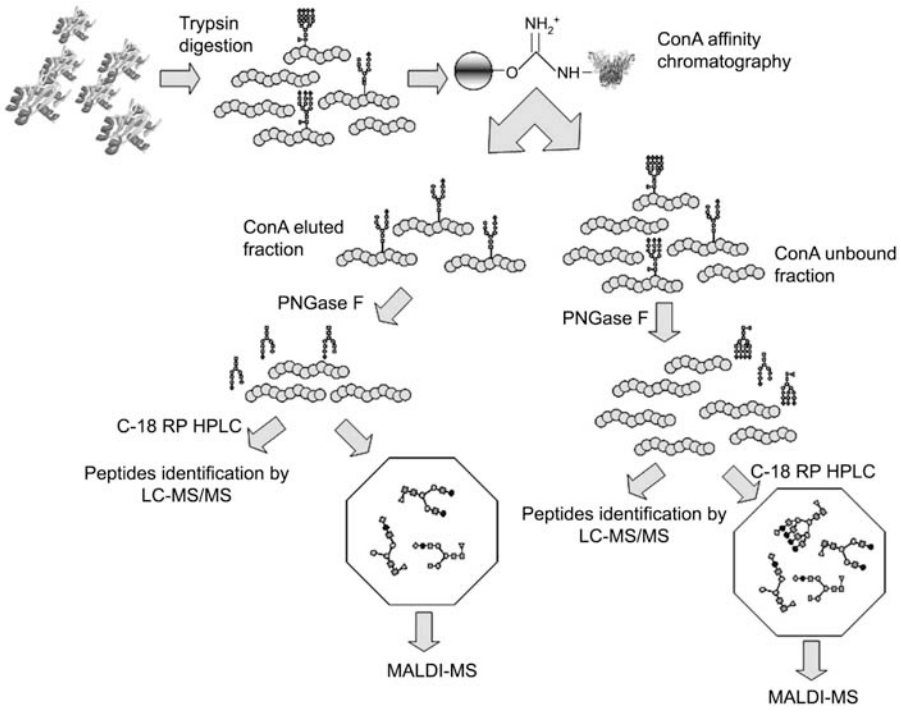


Figure 3 Flow chart showing the strategy to characterize N-linked glycoproteomes in sera from healthy and diseased individuals.

could be recognized in the three glycan profiles, the amounts did not significantly. However, differences in the relative intensities of the most abundant peaks were detected in the samples from patients with disease compared with those from healthy individuals.

Non-glycosylated peptides and peptides containing O-glycans and large complex glycans were not retained on the lectin. These constituted the unbound fraction from the affinity purification. PNGase F was added to this fraction, glycans were separated from peptides by HPLC and analyzed by MALDI-MS. Figure 4A–C shows the negative MALDI spectra of the oligosaccharide unbound fraction of the samples from the diseased and healthy individuals.

The structure of the various glycoforms was inferred on the basis of their mass value and the known biosynthetic pathway. The most common glycans were also selected for PSD fragmentation analysis to confirm oligosaccharide assignments. As an example, Figure 5 shows the PSD spectrum of the glycoform at m/z 2391.10. This glycoform displays the classic fragmentation pattern leading to the reconstruction of the sialylated fucosylated biantenna glycan.

In contrast to the fraction retained with ConA, MALDI-MS analysis of the unbound glycan fraction led to the detection of a plethora of glycoforms, and highlighted differences in the profiling of the three sample types. Glycans from patients with disease were characterized by abnormal glycosylation patterns, with differences in branching, sialylation and fucosylation. These patients showed an excessive expression of peculiar structures, the persistence of incomplete and truncated structures, the accumulation of precursors and the appearance of novel structures. This resulted in a higher degree of heterogeneity compared with healthy patients. The presence of more than 60 different glycoforms was observed in patients with hepatic carcinoma (K-hepatic). The increased size and branching of N-linked glycans created additional sites for terminal sialic acid residues; whose role in promoting metastasis is generally recognized. Myocardial lesions revealed the presence of hyper-fucosylated signals, whereas K-hepatic sera were characterized by hyper-fucosylation and hyper-sialylation.

This study provided an overview of the abnormal glycosylation that occurs in disease. The ConA affinity step was instrumental in depleting commonly occurring glycoforms. Without this step, highly branched, less common glycoforms might have escaped mass spectral analyses. Our data for K-hepatic serum were in perfect agreement with results obtained by Zhao et al. (52) who conducted a similar study on patients with pancreatic cancer.

Selective labeling procedures in PTMs

Identification of post-translationally modified proteins requires direct observation of both the modified peptides and the specific residues involved. Due to the complexity of proteomic samples, identification of modified proteins often relies on the use of ''specific'' antibodies able to recognize individual PTMs (70–75). Once the modified proteins have been labeled with the specific antibody, they can be identified by MS procedures such as peptide mass fingerprinting and/or LC-MS/MS. Frequently, however, the actual modified peptides cannot be identified. Moreover, the occurrence of more than one protein in a specific location of the gel makes identification of PTMs problematic.

Several new methods based on selective labeling of modified peptides have recently been proposed (76–79). The strategy behind these methods consist of the direct selective analysis of labeled peptides with the mass spectrometer without the use of 2D gels and enrichment steps needed to reduce the complexity of the peptide mixture prior to MS analysis. Methods specifically tailored for the qualitative and quantitative determination of PTMs have recently emerged (80–83). The use of stable isotope labeled tags (light and heavy tags) (78, 79) can also address the problem of quantitative determination of PTMs. While the tagging method currently described can target some PTMs more specifically than others, the development of suitable selective chemical labeling strategies that can identify a larger number of PTMs should provide increased confidence in the identification of PTMs.

Reporter ion generating tag strategy (RIGhT)

Several groups have reported on the potential of chemical derivatization in conjunction with ESI-MS for the analysis of synthetic and naturally occurring compounds. Treatment with an appropriate derivatizing agent can improve the ionization of a compound and thus the detection of the analyte by ESI-MS.

Recently, our group proposed a novel strategy for identification of PTMs (79, 84). Our strategy has general applicability and is based on selective labeling of phospho-Ser/Thr residues. We proposed that the derivatizating agent may include a chemical moiety capable of fragmenting in a particular, thus allowing for selective mass spectral analysis of modified peptides through use of MS^n techniques.

This strategy was given the acronym reporter ion generating tag (RIGhT) to indicate a common derivatization method based on the labeling of target residues with reagents capable of generating reporter ions in MS^2/MS^3 experiments. Dansyl chloride (DNS-Cl), a very well-known fluorescent derivatizing reagent commonly used in protein chemistry, was chosen as the reporter tag. Approximately 40 years ago we discovered that fragmentation of dansyl derivatives gave rise to the typical m/z 170 and m/z 234 daughter ions fragments in electron impact experiments (85).

The RIGhT method is based on: 1) selective modification of target residues with DNS-Cl or other available dansyl reagents and, 2) selective detection and identification of labeled peptides by their characteristic fragmentation pathway in MS/MS experiments and the diagnostic m/z 234–170 transition in MS^3 mode. Precursor ions scans (PIS) of the m/z 170 frag-

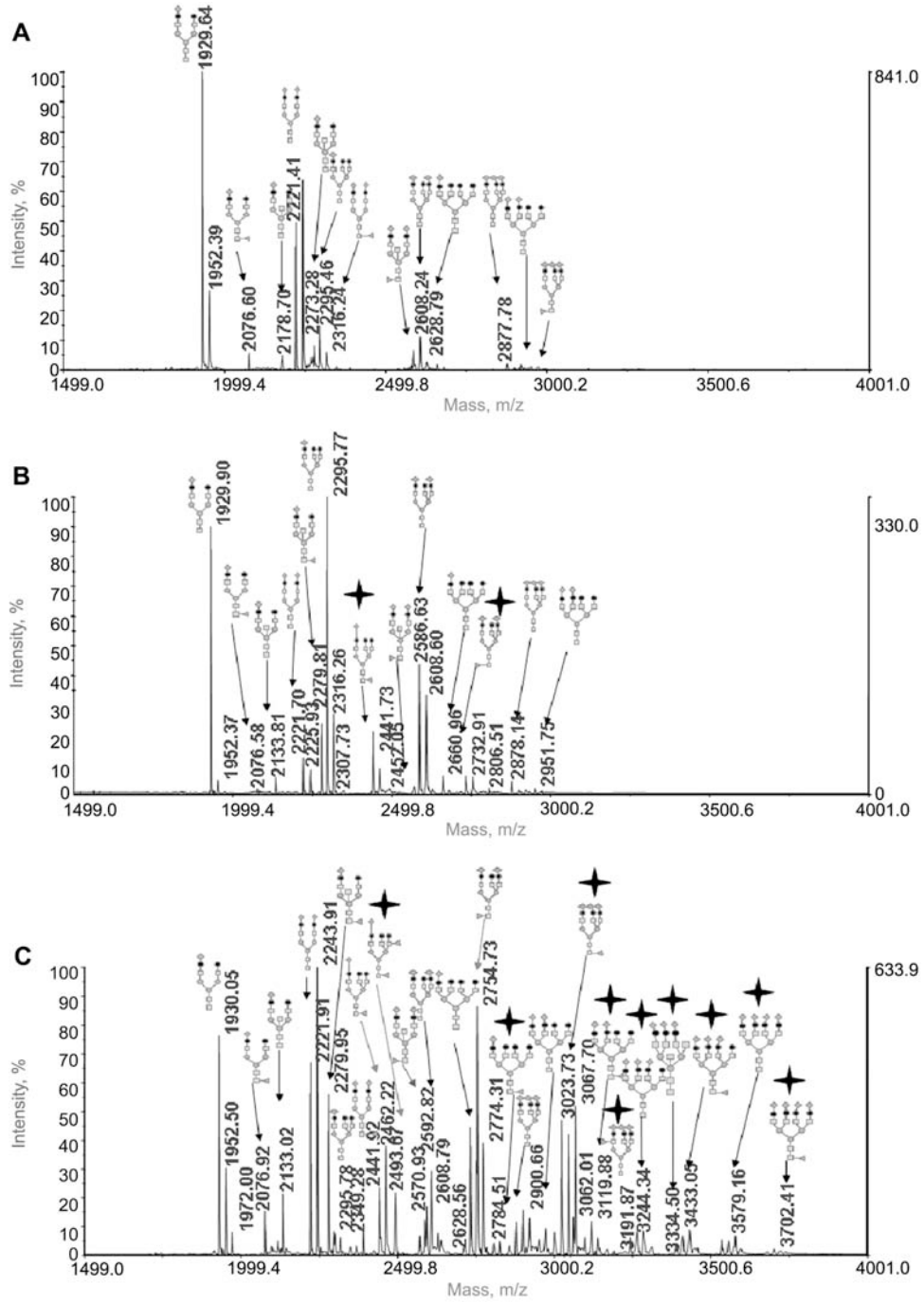


Figure 4 MALDI-MS profiling in negative ion mode of the oligosaccharide extracted from the unbound fraction of ConA affinity chromatography.
(A) Healthy serum, (B) myocardial lesions serum, (C) K-hepatic serum. The most important oligosaccharides are depicted; structures present only in sera from individuals with disease are indicated by a star.

ment can then be combined with MS^2 and MS^3 analysis to specifically detect only the parent ions producing the m/z 234–170 fragmentation (Figure 6).

Application of the RIGhT approach is dependent on the availability of high-performance MS instruments able to perform analysis with high-sensitivity and accuracy. LIT mass spectrometers developed by the combination of triple quadrupole MS with LIT technology is one such device capable of this type of analysis (86). LIT mass spectrometers combine the advantages of very selective scanning of precursor ion with a triple quadrupole along with sensitive MS/MS capabilities of the ion trap (87). This hybrid instrument is suitable for highly sophisticated MS^n analysis.

Phosphorylation

Phosphorylation is a reversible modification that acts as a switch to turn ''on'' or turn ''off'' protein activities or cellular pathways. Phosphorylation acts in a reversible manner by promoting the folding and function of proteins, enzyme activities or substrate specificities,

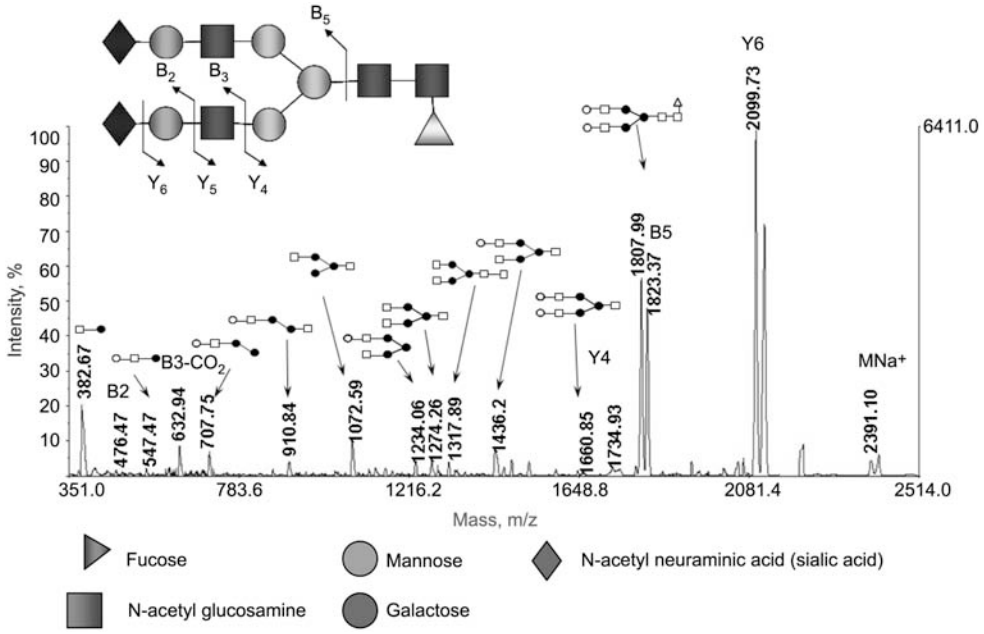


Figure 5 The MNa^+ ion at m/z 2391.10 was selected and submitted to PSD fragmentation.
Y and B ions as well as fragments originated by double cleavages are indicated.

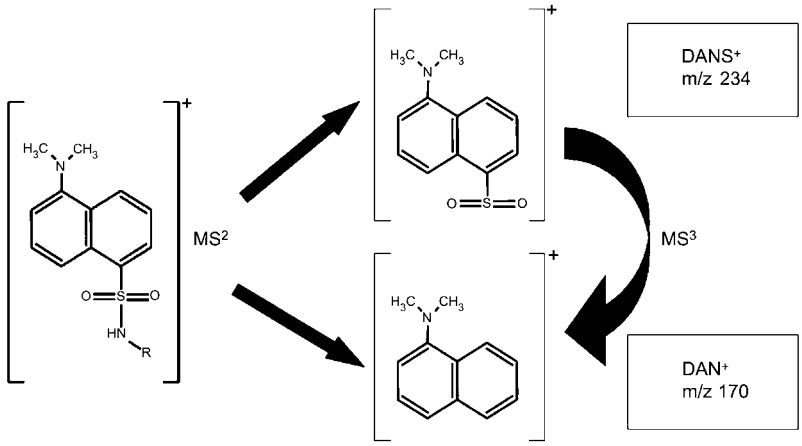


Figure 6 Fragmentation of dansyl derivatives.
The typical m/z 170 and m/z 234 daughter ions fragments in MS/MS experiments as well as the diagnostic m/z 234-170 fragmentation in MS^3 mode are indicated.

regulating protein localization, and complex formation and degradation. A wide range of cellular processes including cell cycle progression, differentiation, development, peptide hormone response, signal transduction, metabolic maintenance and adaptation are all regulated by protein phosphorylation (88). The variety of functions in which phosphoproteins are involved necessitates a wide diversity of phosphorylation. Several amino acids can be phosphorylated by four types of phosphorylation. O-phosphates are the most common class and attach most often to serine-, threonine- and tyrosine-residues. However, O-phosphates can also attach to unusual amino acids such as hydroxyl-proline. N-, S- and acyl-phosphorylation are far less common and occur on histidine and lysine (N-phosphates), cysteine (S-phosphates) and aspartic and glutamic acid residues (acyl-phosphates).

Another characteristic of phosphorylation is the ratio of phosphorylated and non-phosphorylated proteins present in the cell. Protein phosphorylation usually occurs at substoichiometric level in the cell; some residues are always phosphorylated, while others are modified only transiently. Phosphorylation rates and phosphoproteins abundance are very low in signaling pathways, accounting for only 1%–2% of the entire protein content present in the phosphorylated form.

Protein phosphorylation is tightly regulated by a complex set of kinases and phosphatases which are responsible for protein phosphorylation and dephosphorylation, respectively. Kinases and phosphatases are highly regulated, and phosphorylation cycles may take place on a very short timescale. In addition to problems concerning regulation, analysis is compli-

cated by the complexity of phosphorylation patterns and the sheer number of phosphorylation sites. In many cases, phosphorylation effects are combinatorial and multiple sites are phosphorylated (89).

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 100,000 estimated phosphorylation sites in the human proteome. Deciphering the phosphoproteome will be a huge and challenging task due to the different types of phosphorylation generating a variety of phosphoproteins that are not easily analyzed.

Phosphoproteome analysis

An important part of phosphorylation analysis is sample preparation that preserves protein phosphorylation. Detection of phosphorylated proteins within complex mixtures is usually not possible without prior separation of the proteins. Although 2D-gel electrophoresis has been used primarily for separation and quantification of proteins from different systems, methods are also being developed to characterize the phosphoproteome. Phosphoproteins can be visualized directly in 2D gels using phospho-specific stains or Western blotting.

Commercially available phospho-specific stains are less sensitive than methods that use radioisotopes, but the handling of these non-isotopic reagents is more convenient (90). As an example, Pro-Q Diamond, (Molecular Probes) a fluorescent dye, allows detection of 1–20 ng of phosphoproteins 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 principle used for staining, this method is more specific for Ser/Thr phosphorylations.

Antibodies can be used to discriminate between phosphorylation of serine, threonine and tyrosine (91). The specificity and sensitivity of immunostaining is highly dependent on the antibody used. Various antibodies of good specificity toward phospho-tyrosine are available with very little cross-reactivity.

Once phosphorylated, protein can be detected on a gel using either specific staining or immunotechniques. A 2D gel can be prepared and developed using colloidal Coomassie Blue or Silver staining. Following image analysis, phosphorylated protein spots are matched to protein spots on the preparative gel. Proteins of interest can then be excised from the gel, digested and identified by MS. Major advantages of this approach are in the detection of N-, S- and O-phosphorylation and in the possibility of quantitating the modified proteins (92, 93). To measure changes in the state of protein phosphorylation, two distinct proteome samples can be analyzed and immunoblotted, and the intensity of corresponding spots compared.

The main drawback of the method is that indirect identification of phosphorylated proteins by staining methods for direct detection of modified peptides is very unlikely using MS (94, 95). Moreover, when many proteins are present in the same location on the gel, identification of the actual modified protein is not trivial.

Over the years, several strategies have been developed to increase the sensitivity of phosphoproteome analysis that eliminate the need for specific staining and antibody labeling (96, 97).

Recently, MS based methods have emerged as powerful and preferred tools for the analysis of PTMs, including phosphorylation, due to higher sensitivity, selectively, and speed compared with other biochemical techniques. However, despite advances in MS for analysis of phosphorylation, many difficulties remain. First, the generally low phosphorylation stoichiometry of proteins, including phosphopeptides, results in low amounts of protein in the peptide mixtures. Second, the increased hydrophilicity results in reduced retention of phosphopeptides on reversed-phase columns. Finally, there is selective suppression of the ionization/detection efficiency in the presence of large amounts of unphosphorylated peptides when using MS analysis in positive mode. For these reasons, several technologies such as affinity, liquid reverse phase, ion exchange chromatography and capillary electrophoresis, or some combination of these prior to MS analysis, have been used in proteomics to enrich phosphoproteins that might otherwise not be detected. The enrichment step combined with the high-sensitivity of MS technologies provides great potential for characterization of the phosphoproteome (98).

The use of miniaturized immobilized metal affinity chromatography (IMAC), in which phosphopeptides are non-covalently bound to resins that chelate Fe(III) or other trivalent metals, followed by base elution, has proved to be a potentially valuable method for enrichment of phosphopeptides (99–102). With further refinements, this technique may offer the best potential for large-scale phosphorylation analysis.

Another approach to enrich phosphoproteins is based on the use of specific antibodies to immunoprecipate proteins from complex lysates. With this approach, the extraction of phosphoproteins leads to substantial simplification of the protein pattern, and enrichment phosphoproteins present in low concentrations (103).

Several methods for selective enrichment of phosphoproteins and phosphopeptides utilize chemical modification of the phosphate group. This approach does not distinguish between O-glycosylated and phosphorylated analytes, thus requiring additional analysis to confirm phosphorylation. Zhou et al. established a multi-step derivatization method capable of enriching not only Ser/Thr-phosphorylated peptides, but all types of phosphorylated peptides as well (6, 104). Phosphopeptides are bound to sulfidryl-containing compounds via phosphoamidate-bonds and can, therefore, be linked covalently to a solid support with immobilized iodoacetyl-groups. The phosphate-groups are not cleaved off the respective residues, so native phosphopeptides are obtained after elution with trifluoroacetic acid (TFA). A drawback of this

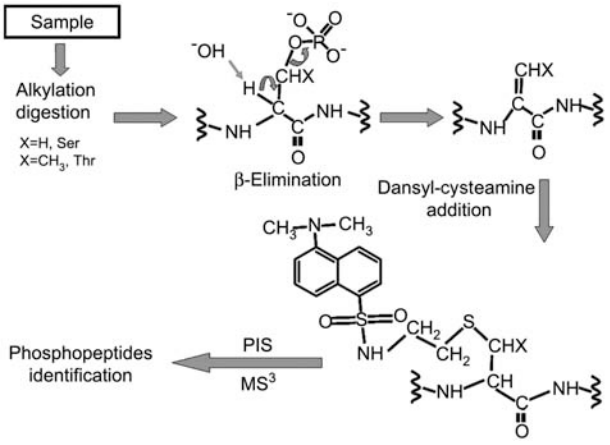


Figure 7 Identification of phosphoSer/Thr by β-elimination of the phosphate group from Ser/Thr residues, selective addition of a dansyl derivative on the newly generated double bond, and detection by precursor ion/MS^2/MS^3 mass spectral analysis.

method is the extensive work is required to quantitatively block all amino- and carboxy-groups within the peptides to prevent intramolecular and intermolecular condensation.

Another chemical tagging reaction is based on the β-elimination of phosphate groups in phospho-Serine/phospho-Threonine (pSer/pTr) residues under alkaline conditions, followed by a Michael-type addition of a thiol compound on the resultant double bond (105).

Dansyl labeling of pSer/pThr

In an attempt to improve and simplify the plethora of methods described, we employed the RIGhT strategy (84) to the selective identification of phosphoSer/Thr peptides. The method consists of (a) β-elimination of the phosphate group from Ser/Thr residues; (b) selective addition of a dansyl derivative on the newly formed double bond, and (c) the selective detection and identification of labeled peptides by precursor ion/MS^2/MS^3 mass spectral analyses (Figure 7).

This approach was first developed using bovine α-casein as a model phosphoprotein. The phosphate moieties from a tryptic α-casein digest were removed from pSer and pThr using barium hydroxide ion-mediated β-elimination (79). The peptide mixture was then analyzed using MALDI-MS which showed ions corresponding to the β-eliminated peptides that were 98 Da lower than expected. However, although DNS-Cl is known to react with primary NH_2 groups, it is unable to give a Michael addition to double bonds. A novel dansyl derivative of cysteamine was, therefore, synthesized following the synthetic route depicted in Figure 8. The free SH group of dansyl-cystamine was then reacted with the β-eliminated peptide mixture via Michael-type addition. The yield of this addition reaction was monitored with MALDI-MS, showing that the extent of the reaction was not quantitative as expected for a typical Michael reaction addition (84).

In a proof-of-principle experiment, a tryptic digest from ten standard proteins was spiked with a minimum amount of an α-casein tryptic mixture modified with dansyl-cysteamine. An aliquot of 1 pmol of the peptide mixture was analyzed by LC-MS/MS with a LIT instrument, combining a PIS (m/z 170) with MS^3 analysis (see above) (Figure 9).

Figure 10 shows a typical reconstructed ion chromatogram for the transition m/z 234/170 in MS^3 mode that is dominated essentially by two signals. The corresponding MS^2 spectra led to reconstruction of the entire sequence of the β-eliminated phosphopeptides 104–119 and 106–119 that carried a dansyl-cysteamine moiety. The modified ion is stable during CID, 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, allowing exact localization of the site of phosphorylation.

However, this approach seems amenable only for monophosphorylated peptides since the diphosphorylated and pentaphosphorylated α-casein tryptic

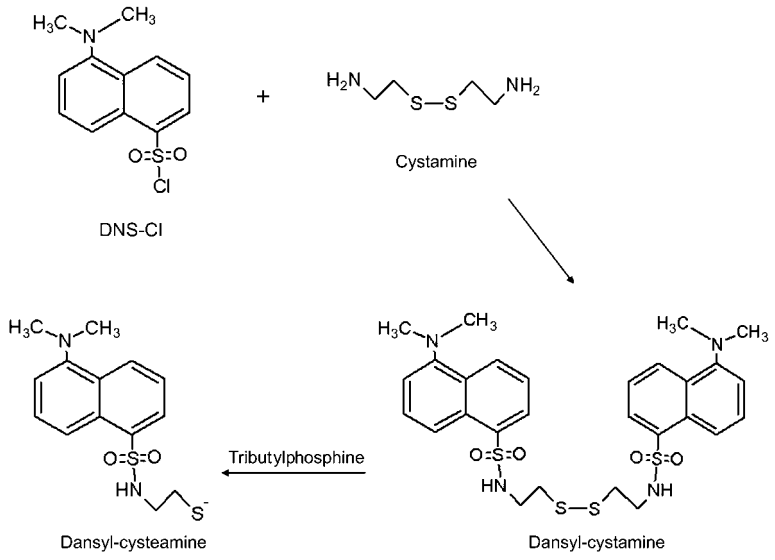


Figure 8 Synthesis of dansyl-cysteamine.

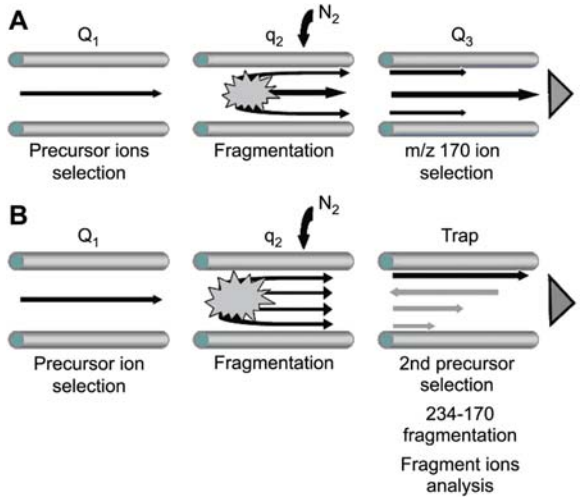


Figure 9 Flow chart showing bidimensional mass spectral analysis.
In PIS mode the Q1 was scanned across the full mass range, and ions are fragmented in the collision cell. The Q3 was set to transmit only the mass of the diagnostic product ion (m/z 170). Therefore, only those precursor ions that originated the ion at m/z 170 by fragmentation were detected (A); these selected precursor ions were then submitted to a combined MS^2/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 detection of the dansyl peptides (B).

peptides could not be detected. This is probably due to steric hindrance and high hydrophobicity of highly modified dansyl peptides. The detection of multiphosphorylated peptides is a difficult challenge. Among the proposed derivatization procedures, only the modification strategy that we have proposed based on the use of dithiothreitol (79) has resulted in mass spectrometric analysis of multiphosphorylated peptides that we have confidence in.

Nitration

Nitration can occur in cells during oxidative stress and inflammation, likely through generation of peroxynitrite ($ONOO^-$) which is a potent oxidizing and nitrating agent. Peroxynitrite is produced by the reaction of superoxide anion in the presence of excess nitric oxide (106, 107). Targets of nitration include nucleic acids, sugars, lipids, alcohols, small organic molecules and aromatic amino acids (108). Sites of nitration in proteins are phenylalanine, tryptophan and tyrosine residues, with the phenolic ring being particularly active toward this modification.

Tyrosine nitration is increasingly recognized as the prevalent, functionally significant post-translational protein modification that serves as an indicator of nitric oxide (NO^-) mediated oxidative inflammatory reactions (109). Tyrosine nitration, in fact, reflects the extent of oxidant production that occurs in both physiological and pathological conditions. Two catalytic mechanisms have been proposed for in vivo nitration. Both mechanisms involve the formation of tyrosyl radicals that can react with nitric dioxide to generate NO_2Tyr. In addition to this principal pathway, tyrosyl radicals can also react with another tyrosyl radical to produce the 3-3′-dityrosine specie which can generate intramolecular cross-links or protein multimerization.

Another possible mechanism of protein nitration in vivo is the reaction of tyrosyl radicals with nitric oxide (NO), resulting in the formation of 3-nitrosotyrosine (108). Apparently, nitrotyrosine is not stable and is oxidized by two electrons to form 3-nitrotyrosine.

Addition of a NO_2 group to the *orto* position of tyrosine decreases the pK_a of the phenolic group approximately three pH units. This gives a net negative charge that is approximately half of that of nitrated tyrosine residues at physiological pH.

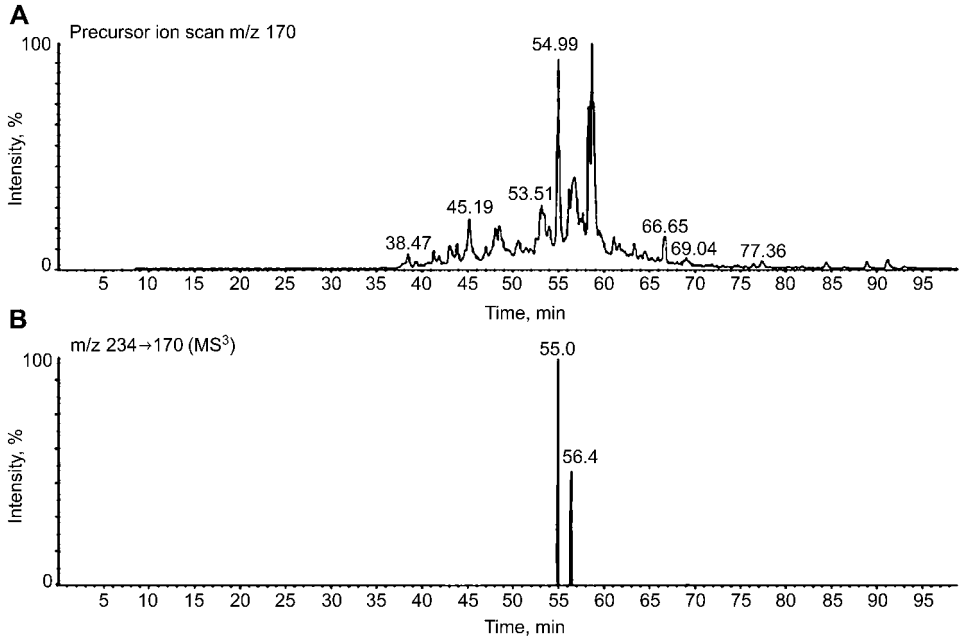


Figure 10 Ion chromatogram for PIS mode analysis (A) and reconstructed ion chromatogram for the transition m/z 234/170 in MS^3 mode (B).
The MS3 TIC showed the presence of only two peaks with a large increase in the signal-to-noise ratio.

Nitration of proteins modulates catalytic activity, cell signaling and cytoskeleton organization. This bulky anionic constituent can induce changes in protein conformation, resulting in the generation of antigen epitopes, altered enzyme catalytic activity, modulation of metabolic pathways and inhibition of tyrosine phosphorylation by protein kinases. NO_2Tyr is considered the most important biomarker for identification of cellular processes associated with RNS formation leading to PTMs of proteins. RNS is the acronym for Reactive Nitrogen Species used to indicate all chemicals derived by nitrogen metabolism. These chemicals have a crucial role in regulation of many biological functions, but if their concentration increases too much, these species might initiate various types of cellular damage. It is well-known that some PTMs due to oxidative stress are linked to the pathogenesis of different diseases (109). In this respect, it is important to find biological markers that can provide information on the localization and physiological effects of oxidative stress. Molecules modified by oxidation, nitration and nitrosilation can act as such cellular biomarkers.

For these reasons many research groups have focused their attention on 3-nitrotyrosine residues, developing various methodologies for identification of NO_2Tyr residues. Protein nitration has been observed in connection with more than 60 human disorders including Alzheimer's disease, atherosclerosis, acute lung disease, amyotrophic lateral sclerosis (ALS) and others. Nitration is detected primarily by immunodetection using antibodies raised against nitrotyrosine (110).

From the limited number of experiments simulating in vivo conditions, it is apparent that tyrosine nitration is a selective process. Not all proteins or all tyrosine residues are targets for nitration (108). However, neither the abundance of a protein nor the abundance of tyrosine residues in a protein can predict nitration. Several factors such as the nitrating agent, protein folding, and the presence of glutamate in the local environment of the tyrosine residue are expected to contribute to nitration of specific tyrosine residues.

Nitroproteome analysis

Detection and quantification of nitroproteins has been accomplished essentially using 2D gel electrophoresis and Western blotting. In particular, a 2D gel containing an aliquot of the protein sample is stained with anti-3-nitrotyrosine antibody. The spots containing modified protein are selectively detected. The remaining portion of the sample can be loaded onto a preparative 2D gel stained either with colloidal Coomassie or Silver stain. Detection of nitrated protein spots is performed using image analysis by comparing the antibody stained and the preparative gels. The corresponding protein can then be identified using conventional mass spectral techniques. Although widely employed, this procedure has a number of drawbacks. Recognition of nitroproteins is done with an antibody which might show some non-specific detection. Moreover, when two or more proteins are present in the same spot, definitative identification of the nitrated component cannot be accomplished. Finally, oftentimes localization of the site of modification within the nitrated protein sequence is impossible unless the modified peptides are detected during the identification procedure.

Other authors have employed methods using liquid chromatography separation combined with electrochemical detection or GC-MS techniques (111, 112). An alternative method to examining the overall amount of 3-nitrotyrosine in proteins is based on the use of electrospray MS. It has been shown by Petersson and co-workers that use of ESI-MS in PIS mode for detection of nitrated peptides in complex mixtures can be accomplished by taking advantage of characteristic fragment ions of modified peptides (113). The immonium ion of nitrotyrosine, in fact, provided unequivocal evidence for the presence of this modification, suggesting that mass spectral analyses of intact peptides might be a suitable analytical alternative.

The spectrum of nitrated peptides identified by use of MALDI shows a specific and unique pattern of molecular ions that can be used to identify this modification (113). The mass signal of the modified peptide occurs 45 Da higher than the unmodified specie due to substitution of an aromatic hydrogen with NO_2. This signal is usually accompanied by satellite ions that are 16, 30 and 32 Da lower in mass. This unique molecular ion pattern generated by photodecomposition during MALDI analysis provides a unique signature for peptides containing 3-nitrotyrosine. This fragmentation of nitrated peptides observed with MALDI-MS reduces the abundance of signals for nitrated peptides. This results in failure to detect modified peptides in very complex mixtures. It has been shown by Petersson and co-workers (113) that ESI-MS analysis of nitrated peptides prevents photodecomposition and allows the nitrated peptides to be specifically detected in PIS mode using the immonium ion of nitrotyrosine.

Dansyl labeling for the detection of 3-nitrotyrosine residues

We extended the RIGhT strategy to the selective isolation and identification of o-nitrotyrosine-containing proteins (114). The philosophy of the method consisted of the specific reduction of 3-nitroTyr to the corresponding aminotyrosine (NH_2-Tyr). The amino group of this residue is endowed with a particular low pK_a value, thus allowing selective labeling with DNS-Cl at variance with aliphatic (10.4–11.1) and N-terminal amino groups (6.8–8.0), which are largely protonated.

The methodology was first tested on bovine serum albumin (BSA) nitrated in vitro as a model protein and then applied to more complex matrices. BSA was nitrated in vitro with tetranitromethane (TNM) and the TNM-induced modifications were determined with MALDI MS. Spectral analyses led to 85% of BSA sequence coverage. More importantly, however, 95% of tyrosine residues could be verified. Two mass signals recorded at m/z 972.5 and m/z 1524.8 occurred 45 Da higher than the signals corresponding to the peptides 137–143 and 396–409, respectively. This

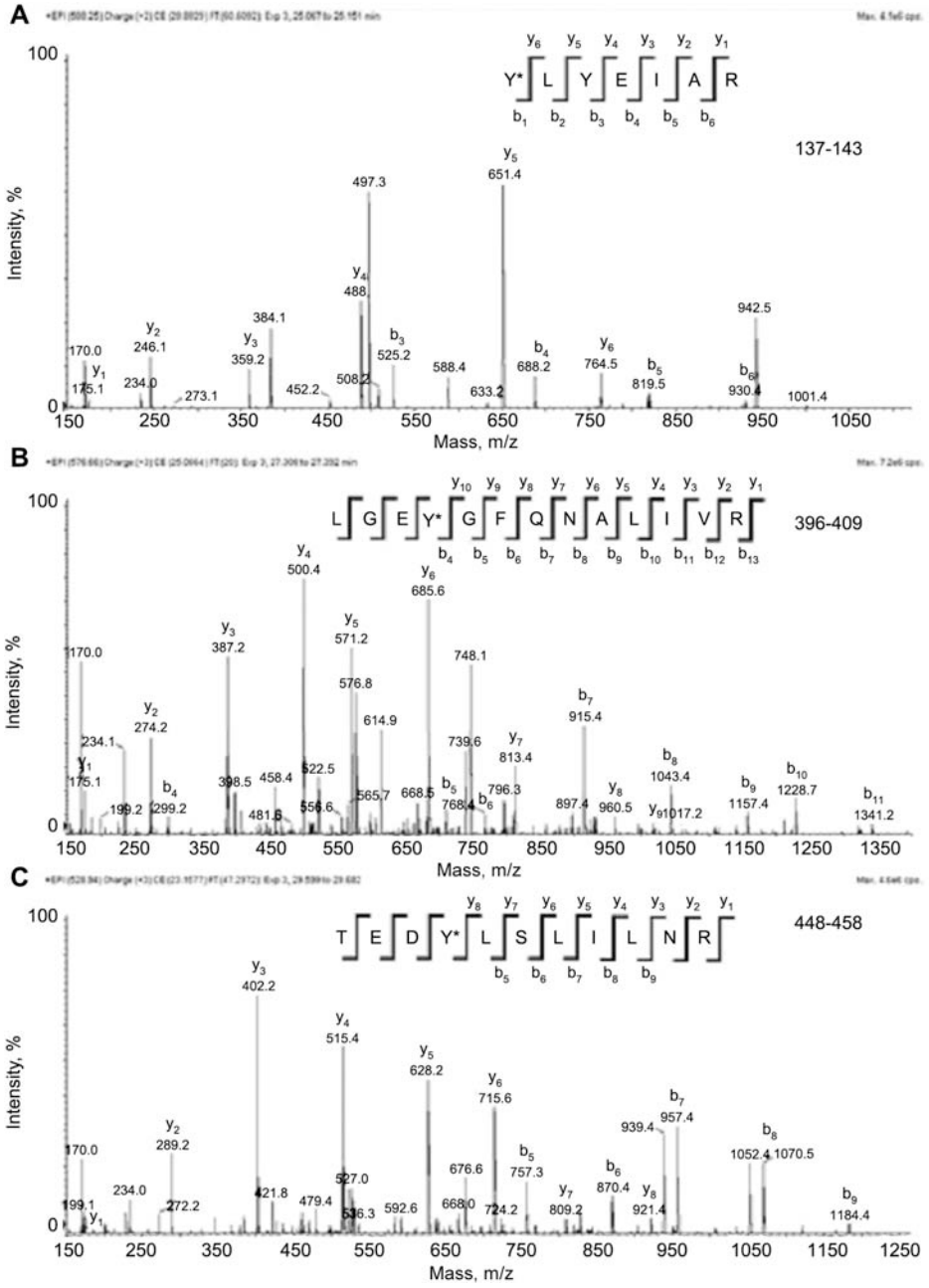


Figure 11 MS/MS spectra of nitrated BSA peptides 137–143 (A), 396–409 (B) and 448–458 (C).
Fragment ions are labeled as ''y'' or ''b'' according to the common peptide MS/MS fragmentation rules. Interpretation of the spectra led to the localization of the original nitration sites and to the determination of the peptide sequence. Unlabeled signals originated from secondary fragmentation pathways. Dansylated Tyr residues located at the level of Tyr137, Tyr399, and Tyr451, are highlighted as Y* in the sequence.

suggests that they originated from the nitrated derivatives of these fragments. However, signals corresponding to the unmodified fragments were also detected, indicating that the nitration reaction was not quantitative.

Conversion of nitropeptides into their o-dansylamino derivatives was accomplished by treatment with $Na_2S_2O_4$. Aminotyrosine residues were then selectively labeled with DNS-Cl following the procedure reported by Cirulli et al. (115). The reaction was carried out at pH 5.0, exploiting the different pK_a-value of the o-aminotyrosine (4.7). MALDI-MS analysis showed the presence of two new signals occurring 233 Da higher than the o-aminotyrosine-containing peptides, corresponding to the expected dansyl derivatives. Signals corresponding to the unmodified o-aminotyrosine peptides indicated that the extent of dansyl modification was about 60%. No modification of the N-terminus or Lys residues was detected.

The exact location of the original nitro groups was assessed by LC-MS/MS analysis of the peptide mixture using the experimental procedure described above (101, 103). LC-MS/MS full scan analysis showed a highly complex mixture of peaks that prevented identification of the modified peptides. The total ion current (TIC) trace for the m/z 170 PIS dis-

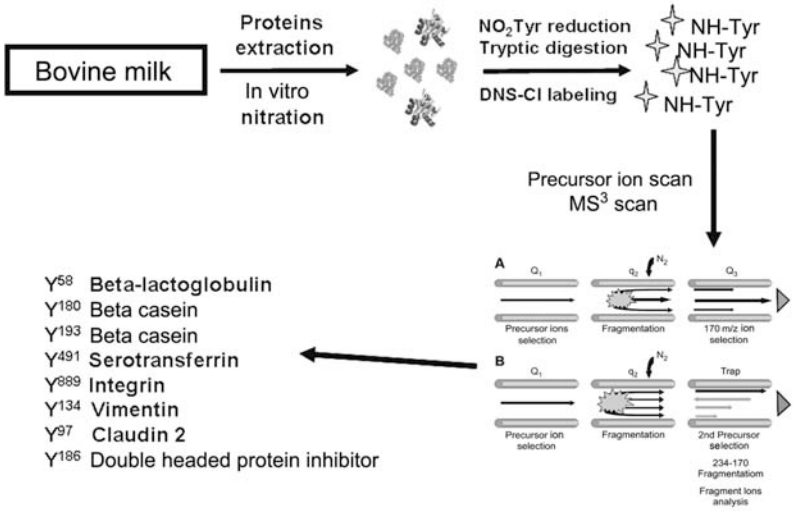


Figure 12 Flow chart of the analytical strategy for detecting labeled nitroproteins in bovine milk protein extract.

played a simplierd chromatogram with significantly fewer number of signals. However, too many peaks were still present and it was very unlikely they could be related to dansyl derivatized peptides only. Finally, once the m/z 170 precursor ion and the m/z 234/170 transition scan selectivity criteria were applied, the MS^3 TIC showed the presence of only three peaks with a large increase in the signal-to-noise ratio. The corresponding MS^2 fragmentation spectra enabled the determination of the entire sequence of these species and allowed definitive identification of the peptides. Of greater interest was the fact that the dansylated Tyr residues were located at the level of Tyr137, Tyr399 and Tyr451. The allowed for exact identification of the original nitration sites (Figure 11). It should be noted that LC-MS/MS analysis revealed the occurrence of a further nitration site at level of Tyr451 within the peptide 448–458, which had escaped previous MALDI analysis.

The MS/MS spectra of o-dansylamino-Tyr-containing peptides showed additional peculiar features. A stable fragment signal was always observed at m/z 384 corresponding to the immonium ion of o-dansylamino-Tyr. This could then be used either as an alternative diagnostic fragment in the PIS, or to confirm the presence of an o-dansylamino-Tyr residue in the peptide sequence.

The feasibility of the MS procedure to identify *o*-nitrotyrosine peptides in proteomics was evaluated by mixing 100 μg of a mixture of BSA and nitrated BSA with 10 mg of the entire cellular extract from *Escherichia coli*. The total protein extract was then submitted to the procedure described above. The selection based on the MS^3 scan removed a large number of false positives leading to a simple ion chromatogram that was dominated essentially by three intense signals. The MS^2 fragmentation spectra of these signals confirmed the occurrence of the same three BSA nitrated peptides detected previously.

The proposed strategy was finally employed to identify o-nitrotyrosine residues in a more complex sample, bovine milk (114). Bovine milk was nitrated in vitro with TNM and an aliquot of milk proteins were fractionated using SDS-PAGE and submitted to Western blot analysis using antibody against o-nitrotyrosine to confirm the extent of nitration. The presence of nitrated proteins was detected essentially in the regions between 20–35 and 45–60 kDa. The entire milk protein extract was dissolved in denaturizing buffer and cysteine alkylation and reduction of nitro groups were performed LC-MS/MS analysis using the double selectivity criteria led to the identification of nitropeptides from the high-abundant milk proteins, α-casein and α-lactoglobulin. In addition, it was able to assess the nitration sites occurring in proteins present in low concentration, such as vimentin and claudin 2 (Figure 12).

Conclusions

At the turn of the last century, the international scientific community witnessed a tremendous burst of interest in the area of expression proteomics. There was an enormous increase in numbers of papers being published, new journals being launched and large amounts of money invested by the pharmaceutical industry in search of new biomarkers of disease and therapeutic targets. In the last few years, interest has subsided slightly due to the lack of any new effective biomarkers being discovered. Also, new applications in proteomics have begun to emerge. It is easy to foresee that in the near future, most of proteomic research will be addressed toward two main goals; the investigation of cellular mechanisms through isolation of functional complexes, and qualitative and quantitative analysis of PTMs.

As discussed in this review, assessment of substoichiometric and transient PTMs on a proteomic scale is still a challenge. Existing procedures are not endowed with general applicability. The successful identification of PTMs is still dependent on a number of factors including the nature and the extent of the modifications, the amount of sample available, and

other variables. Despite advances in technology, the proteomic approach to PTMs still seem far from being satisfactory and productive.

The recent trend shown by research in this field suggests that scientists are moving along two main paths that might also be combined. One path involves particular attention to areas related to sample preparation and separation technologies, with particular emphasis on enrichment procedures. The goal is to develop selective methodologies able to increase the amount of modified samples when compared to unmodified ones.

The other path involves efforts made toward addressing the utilization of advanced analytical chemistry strategies focused on the development of new instrumentation and technologies in the field of MS. The introduction of sophisticated instrumentation, like the Orbitrap, or totally new concepts of ion separation such as the ion mobility MS might open up ways to utilize ''bidimensional'' or ''multidimensional'' mass spectral analysis able to identify modified peptide ions in very complex mixtures.

The integration of innovative enrichment procedures with these new mass spectrometric techniques should enable a new generation of proteomic experiments to be performed, thus providing a more general approach to analysis of PTMs. In this respect, specific chemical labeling of modified residues allowing for both selective enrichment with affinity tags, and sophisticated mass spectral experiments using reporter tag fragmentation and quantitative analysis by labeling with isotopes, will play a pivotal role in the next generation of proteomics.

References

1. Godovac-Zimmermann J, Brown LR. Perspectives for mass spectrometry and functional proteomics. Mass Spectrom Rev 2001;20:1–57.
2. Patterson SD, Aebersold RH. Proteomics: the first decade and beyond. Nat Genet 2003;33:311–23.
3. Taylor SW, Fahy E, Ghosh SS. Global organellar proteomics. Trends Biotechnol 2003;21:82–8.
4. Mann M, Hendrickson RC, Pandey A. Analysis of proteins and proteomes by mass spectrometry. Annu Rev Biochem 2001;10:437–73.
5. Banks RE, Dunn MJ, Hochstrasser DF, Sanchez JC, Blackstock W, Pappin DJ, et al. Proteomics: new perspectives, new biomedical opportunities. Lancet 2000;356:1749–56.
6. Oda Y, Nagasu T, Chait BT. Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. Nat Biotechnol 2001;19:379–82.
7. Riederer BM. Non-covalent and covalent protein labeling in two-dimensional gel electrophoresis. J Proteomics 2008;71:231–44.
8. Gesellchen F, Bertinetti O, Herberg FW. Analysis of post-translational modifications exemplified using protein kinase A. Biochim Biophys Acta 2006;1764:1788–800.
9. Han KK, Martinage A. Post-translational chemical modifications of proteins – III. Current developments in analytical procedures of identification and quantitation of post-translational chemically modified amino acid(s) and its derivatives. Int J Biochem 1993;25:957–70.
10. Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal Chem 1988;60:2299–301.
11. Beavis RC, Chait BT. Rapid, sensitive analysis of protein mixtures by mass spectrometry. Proc Natl Acad Sci USA 1990;87:6873–7.
12. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Electrospray ionization for mass spectrometry of large biomolecules. Science 1989;246:64–71.
13. Chait BT, Kent SB. Weighing naked proteins: practical, high-accuracy mass measurement of peptides and proteins. Science 1992;257:1885–94.
14. Lee TD, Moore RE, Young MK. Introducing samples directly into electrospray ionization mass spectrometers using a nanospray interface. Curr Protoc Protein Sci 2001;5:unit 16.9.
15. Dal Piaz F, De Leo M, Braca A, De Simone F, Morelli I, De Tommasi N. Electrospray ionization mass spectrometry for identification and structural characterization of pregnane glycosides. Rapid Commun Mass Spectrom 2005;19:1041–52.
16. Schwartz JC, Jardine I. Quadrupole ion trap mass spectrometry. Methods Enzymol 1996;270:552–86.
17. Macek B, Waanders LF, Olsen JV, Mann M. Top-down protein sequencing and MS3 on a hybrid linear quadrupole ion trap-orbitrap mass spectrometer. Mol Cell Proteomics 2006;5:949–58.
18. Hager JW. A new linear ion trap mass spectrometer. Rapid Commun Mass Spectrom 2002;16:512–26.
19. Hager JW, LeBlanc JC. High-performance liquid chromatography-tandem mass spectrometry with a new quadrupole/linear ion trap instrument. J Chromatogr A 2003;1020:3–9.
20. Roepstorff P, Fohlman J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. Biomed Mass Spectrom 1984;11:601.
21. Hanrieder J, Nyakas A, Naessén T, Bergquist J. Proteomic analysis of human follicular fluid using an alternative bottom-up approach. J Proteome Res 2008;7:443–9.
22. Monti M, Orrù S, Pagnozzi D, Pucci P. Interaction proteomics. Biosci Rep 2005;25:45–56.
23. Monti M, Orrù S, Pagnozzi D, Pucci P. Functional proteomics. Clin Chim Acta 2005;357:140–50.
24. Zhang K, Wrzesinski K, Fey SJ, Mose Larsen P, Zhang X, Roepstorff P. Assessing CMT cell line stability by two dimensional polyacrylamide gel electrophoresis and mass spectrometry based proteome analysis. J Proteomics 2008;71:160–7.
25. Li Z, Zhao X, Bai S, Wang Z, Chen L, Wei Y, et al. Proteomic identification of cyclophilin A as a potential prognostic factor and therapeutic target in endometrial carcinoma. Mol Cell Proteomics 2008. [Epub ahead of print].
26. Pucci P, Carestia C, Fioretti G, Mastrobuoni AM, Pagano L. Protein fingerprint by fast atom bombardment mass spectrometry: characterization of normal and variant human haemoglobins. Biochem Biophys Res Commun 1985;130:84–90.
27. Mead JA, Shadforth IP, Bessant C. Public proteomic MS repositories and pipelines: available tools and biological applications. Proteomics 2007;7:2769–86.
28. Macek B, Waanders LF, Olsen JV, Mann M. Top-down protein sequencing and MS3 on a hybrid linear quadrupole ion trap-orbitrap mass spectrometer. Mol Cell Proteomics 2006;5:949–58.
29. Gafken PR, Lampe PD. Methodologies for characterizing phosphoproteins by mass spectrometry. Cell Commun Adhes 2006;13:249–62.
30. Walsh TC. Posttranslational modification of proteins: expanding nature's Inventory. Roberts and Company Publishers, 2006:490 pp.

31. Walsh CT, Garneau-Tsodikova S, Gatto GJ Jr. Protein post-translational modifications: the chemistry of proteome diversifications. Angew Chem Int Ed Engl 2005; 44:7342–72.
32. Spiro RG. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. Glycobiology 2002;12:43R–56R.
33. Qiu R, Regnier FE. Comparative glycoproteomics of N-linked complex-type glycoforms containing sialic acid in human serum. Anal Chem 2005;77:7225–31.
34. Rademacher TW, Parekh RB, Dwek RA. Glycobiology. Annu Rev Biochem 1988;57:785–838.
35. Marshall RD. The nature and metabolism of the carbohydrate-peptide linkages of glycoproteins. Biochem Soc Symp 1974;17–26.
36. Kornfeld R, Kornfeld S. Assembly of asparagine-linked oligosaccharides. Annu Rev Biochem 1985;54:631–64.
37. Parodi AJ. Role of N-oligosaccharide endoplasmic reticulum processing reactions in glycoprotein folding and degradation. Biochem J 2000;348:1–13.
38. Dell A, Morris HR. Glycoprotein structure determination by mass spectrometry. Science 2001;291:2351–6.
39. Perez-Vilar J, Hidalgo J, Velasco A. Presence of terminal N-acetylgalactosamine residues in subregions of the endoplasmic reticulum is influenced by cell differentiation in culture. J Biol Chem 1991;266:23967–76.
40. Spiro RG. Glucose residues as key determinants in the biosynthesis and quality control of glycoproteins with N-linked oligosaccharides. J Biol Chem 2000;275:35657–60.
41. Helenius A. How N-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum. Mol Biol Cell 1994;5:253–65.
42. Trombetta ES. The contribution of N-glycans and their processing in the endoplasmic reticulum to glycoprotein biosynthesis. Glycobiology 2003;13:77R–91R.
43. Varki A. Biological roles of oligosaccharides: all of the theories are correct. Glycobiology 1993;3:97–130.
44. Gagneux P, Varki A. Evolutionary considerations in relating oligosaccharide diversity to biological function. Glycobiology 1999;9:747–55.
45. Karlsson KA. Microbial recognition of target-cell glycoconjugates. Curr Opin Struct Biol 1995;5:622–35.
46. Dell A, Morris HR. Glycoprotein structure determination by mass spectrometry. Science 2001;291:2351–6.
47. Bretthauer RK. Characterization of O-linked saccharides on glycoproteins. Methods Mol Biol 2007;389:107–18.
48. Xie Y, Liu J, Zhang J, Hedrick JL, Lebrilla CB. Method for the comparative glycomic analyses of O-linked, mucin-type oligosaccharides. Anal Chem 2004;76:5186–97.
49. Freeze HH, Kranz C. Endoglycosidase and glycoamidase release of N-linked oligosaccharides. Curr Protoc Protein Sci 2006;12.
50. Zhang H, Li XJ, Martin DB, Aebersold R. Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. Nat Biotechnol 2003;21:660–6.
51. Yang Z, Hancock WS. Approach to the comprehensive analysis of glycoproteins isolated from human serum using a multi-lectin affinity column. J Chromatogr A 2004;1053:79–88.
52. Zhao J, Simeone DM, Heidt D, Anderson MA, Lubman DM, et al. Comparative serum glycoproteomics using lectin selected sialic acid glycoproteins with mass spectrometric analysis: application to pancreatic cancer serum. J Proteome Res 2006;5:1792–802.
53. An HJ, Miyamoto S, Lancaster KS, Kirmiz C, Li B, Lam KS, et al. Profiling of glycans in serum for the discovery of potential biomarkers for ovarian cancer. J Proteome Res 2006;5:1626–35.
54. Durham M, Regnier FE. Targeted glycoproteomics: serial lectin affinity chromatography in the selection of O-glycosylation sites on proteins from the human blood proteome. J Chromatogr A 2006;1132:165–73.
55. Dalpathado DS, Desaire H. Glycopeptide analysis by mass spectrometry. Analyst 2008;133:731–8.
56. Ciardiello MA, D'Avino R, Amoresano A, Tuppo L, Carpentieri A, Carratore V, et al. The peculiar structural features of kiwi fruit pectin methylesterase: amino acid sequence, oligosaccharides structure, and modeling of the interaction with its natural proteinaceous inhibitor. Proteins 2008;71:195–206.
57. Amoresano A, Siciliano R, Orrù S, Napoleoni R, Altarocca V, De Luca E, et al. Structural characterisation of human recombinant glycohormones follitropin, lutropin and choriogonadotropin expressed in Chinese hamster ovary cells. Eur J Biochem 1996;242:608–18.
58. De Lorenzo C, Cozzolino R, Carpentieri A, Pucci P, Laccetti P, D'Alessio G. Biological properties of a human compact anti-ErbB2 antibody. Carcinogenesis 2005;26: 1890–5.
59. Lityńska A, Pocheć E, Hoja-Lukowicz D, Kremser E, Laidler P, Amoresano A, et al. The structure of the oligosaccharides of alpha3beta1 integrin from human ureter epithelium (HCV29) cell line. Acta Biochim Pol 2002; 49:491–500.
60. Inforzato A, Peri G, Doni A, Garlanda C, Mantovani A, Bastone A, et al. Structure and function of the long pentraxin PTX3 glycosidic moiety: fine-tuning of the interaction with C1q and complement activation. Biochemistry 2006;45:11540–51.
61. Zaia J. Mass spectrometry of oligosaccharides. Mass Spectrom Rev 2004;23:161–227.
62. Harvey DJ. Matrix-assisted laser desorption/ionisation mass spectrometry of oligosaccharides and glycoconjugates. J Chromatogr A 1996;720:429–46.
63. Domon B, Costello CE. A systematic nomenclature for carbohydrate fragmentations in FABMS/MS of glycoconjugates. Glycoconjugate J 1988;5:397–409.
64. Sandra K, Devreese B, Van Beeumen J, Stals I, Claeyssens M. The Q-Trap mass spectrometer, a novel tool in the study of protein glycosylation. J Am Soc Mass Spectrom 2004;15:413–23.
65. Meldrum DR. Tumor necrosis factor in the heart. Am J Physiol 1998;274:577–595.
66. Renkonen J, Tynninen O, Häyry P, Paavonen T, Renkonen R. Glycosylation might provide endothelial zip codes for organ-specific leukocyte traffic into inflammatory sites. Am J Pathol 2002;161:543–50.
67. Dube DH. Bertozzi CR. Glycans in cancer and inflammation – potential for therapeutics and diagnostics. Nat Rev Drug Discov 2005;4:477–88.
68. Unkenborg J, Pilch BJ, Podtelejnikov AV, Wiśniewski JR. Screening for N-glycosylated proteins by liquid chromatography mass spectrometry. Proteomics 2004;4: 454–65.
69. Nikov G, Bhat V, Wishnok JS, Tannenbaum SR. Analysis of nitrated proteins by nitrotyrosine-specific affinity probes and mass spectrometry. Anal Biochem 2003; 320:214–22.
70. Deonarain MP. Recombinant antibodies for cancer therapy. Expert Opin Biol Ther 2008;8:1123–41.
71. Go EP, Irungu J, Zhang Y, Dalpathado DS, Liao HX, Sutherland LL, et al. Glycosylation site-specific analysis of HIV envelope proteins (JR-FL and CON-S) reveals major differences in glycosylation site occupancy, glycoform profiles, and antigenic epitopes' accessibility. J Proteome Res 2008;7:1660–74.
72. Kumar GK, Prabhakar NR. Post-translational modification of proteins during intermittent hypoxia. Respir Physiol Neurobiol 2008. [Epub ahead of print].

73. Endoh H, Ishibashi Y, Yamaki E, Yoshida T, Yajima T, Kimura H, et al. Immunohistochemical analysis of phosphorylated epidermal growth factor receptor might provide a surrogate marker of EGFR mutation. Lung Cancer 2008. [Epub ahead of print].
74. Sabetkar M, Low SY, Bradley NJ, Jacobs M, Naseem KM, Richard Bruckdorfer K. The nitration of platelet vasodilator stimulated phosphoprotein following exposure to low concentrations of hydrogen peroxide. Platelets 2008;19:282–92.
75. Larrainzar E, Urarte E, Auzmendi I, Ariz I, Arrese-Igor C, González EM, et al. Use of recombinant iron-superoxide dismutase as a marker of nitrative stress. Methods Enzymol 2008;437:605–18.
76. Baty JW, Hampton MB, Winterbourn CC. Proteomic detection of hydrogen peroxide-sensitive thiol proteins in Jurkat cells. Biochem J 2005;390:791–2.
77. Härtig W, Seeger J, Naumann T, Brauer K, Brückner G. Selective in vivo fluorescence labelling of cholinergic neurons containing p75(NTR) in the rat basal forebrain. Brain Res 1998;808:155–65.
78. Julka S, Regnier F. Quantification in proteomics through stable isotope coding: a review. J Proteome Res 2004; 3:350–63.
79. Amoresano A, Marino G, Cirulli C, Quemeneur E. Mapping phosphorylation sites: a new strategy based on the use of isotopically labelled DTT and mass spectrometry. Eur J Mass Spectrom 2004;10:401–12.
80. Steen H, Kuster B, Fernandez M, Pandey A, Mann M. Tyrosine phosphorylation mapping of the epidermal growth factor receptor signaling pathway. J Biol Chem 2002;277:1031–9.
81. Kalume DE, Molina H, Pandey A. Tackling the phosphoproteome: tools and strategies. Curr Opin Chem Biol 2003;7:64–9.
82. Mann M, Jensen ON. Proteomic analysis of post-translational modifications. Nat Biotechnol 2003;21:255–61.
83. Jensen ON. Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. Curr Opin Chem Biol 2004;8:33–41.
84. Amoresano A, Monti G, Cirulli C, Marino G. Selective detection and identification of phosphopeptides by dansyl MS/MS/MS fragmentation. Rapid Commun Mass Spectrom 2006;20:1400–4.
85. Marino G, Buonocore V. Mass-spectrometric identification of 1-dimethylaminoaphthalene-5-sulphonyl-amino acids. Biochem J 1968;110:603–4.
86. LeBlanc JC, Hager JW, Ilisiu AM, 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;3:859–69.
87. Hager JW, Yves Le Blanc JC. Product ion scanning using a Q-q-Q linear ion trap (Q TRAP) mass spectrometer. Rapid Commun Mass Spectrom 2003;17:1056–64.
88. Ben-Levy R, Leighton IA, Doza YN, Attwood P, Morrice N, Marshall CJ, et al. Identification of novel phosphorylation sites required for activation of MAPKAP kinase-2. EMBO J 1995;14:5920–30.
89. Bouras T, Southey MC, Venter DJ. Overexpression of the steroid receptor coactivator AIB1 in breast cancer correlates with the absence of estrogen and progesterone receptors and positivity for p53 and HER2/neu. Cancer Res 2001;61:903–7.
90. Meimoun P, Ambard-Bretteville F, Colas-des Francs-Small C, Valot B, Vidal J. Analysis of plant phosphoproteins. Anal Biochem 2007;371:238–46.
91. McLachlin DT, Chait BT. Analysis of phosphorylated proteins and peptides by mass spectrometry. Curr Opin Chem Biol 2001;5:591–602.
92. Lewandrowski U, Sickmann A, Cesaro L, Brunati AM, Toninello A, Salvi M. Identification of new tyrosine phosphorylated proteins in rat brain mitochondria. FEBS Lett 2008;582:1104–10.
93. Paradela A, Albar JP. Advances in the analysis of protein phosphorylation. J Proteome Res 2008;7:1809–18.
94. Pandey A, Andersen JS, Mann M. Use of mass spectrometry to study signaling pathways. Sci STKE 2000;2000:PL1.
95. 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;15:1593–9.
96. Idris N, Carothers Carraway CA, Carraway KL. Differential localization of ErbB2 in different tissues of the rat female reproductive tract: implications for the use of specific antibodies for ErbB2 analysis. J Cell Physiol 2001;189:162–70.
97. Bendt AK, Burkovski A, Schaffer S, Bott M, Farwick M, Hermann T. Towards a phosphoproteome map of Corynebacterium glutamicum. Proteomics 2003;3:1637–46.
98. Gaberc-Porekar V, Menart V. Perspectives of immobilized-metal affinity chromatography. J Biochem Biophys Methods 2001;49:335–60.
99. Andersson L, Porath J. ''Isolation of phosphoproteins by immobilized metal (Fe3+) affinity chromatography''. Anal Biochem 1986;154:250–4.
100. Zhou H, Ye M, Dong J, Han G, Jiang X, Wu R, et al. Specific phosphopeptide enrichment with immobilized titanium ion affinity chromatography adsorbent for phosphoproteome analysis. J Proteome Res 2008. [Epub ahead of print].
101. Wilson-Grady JT, Villén J, Gygi SP. Phosphoproteome analysis of fission yeast. J Proteome Res 2008;7: 1088–97.
102. Zhang X, Ye J, Jensen ON, Roepstorff P. Highly efficient phosphopeptide enrichment by calcium phosphate precipitation combined with subsequent IMAC enrichment. Mol Cell Proteomics 2007;6:2032–42.
103. Oda Y, Nagasu T, Chait BT. Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. Nat Biotechnol 2001;19:379–82.
104. Zhou H, Watts JD, Aebersold R. A systematic approach to the analysis of protein phosphorylation. Nat Biotechnol 2001;19:375–8.
105. Mattila K, Siltainsuu J, Balaspiri L, Ora M, Lönnberg H. Derivatization of phosphopeptides with mercapto- and amino-functionalized conjugate groups by phosphate elimination and subsequent Michael addition. Org Biomol Chem 2005;3:3039–44.
106. Beckman JS, Koppenol WH. Nitric oxide, superoxide, and peroxynitrite: the good, the bad, and ugly. Am J Physiol 1996;271:1424–37.
107. Schopfer FJ, Baker PR, Freeman BA. NO-dependent protein nitration: a cell signaling event or an oxidative inflammatory response? Trends Biochem Sci 2003; 28:646–54.
108. Ischiropoulos H. Biological tyrosine nitration: a pathophysiological function of nitric oxide and reactive oxygen species. Arch Biochem Biophys 1998;356:1–11.
109. Greenacre SA, Ischiropoulos H. Tyrosine nitration: localisation, quantification, consequences for protein function and signal transduction. Free Radic Res 2001;34:541–81.
110. Castegna A, Thongboonkerd V, Klein JB, Lynn B, Markesbery WR, Butterfield DA. Proteomic identification of nitrated proteins in Alzheimer's disease brain. J Neurochem 2003;85:1394–401.
111. Aulak KS, Koeck T, Crabb JW, Stuehr DJ. Dynamics of protein nitration in cells and mitochondria. Am J Physiol Heart Circ Physiol 2004;286:30–38.

112. Nikov G, Bhat V, Wishnok JS, Tannenbaum SR. Analysis of nitrated proteins by nitrotyrosine-specific affinity probes and mass spectrometry. Anal Biochem 2003;320:214–22.
113. Petersson AS, Steen H, Kalume DE, Caidahl K, Roepstorff P. Investigation of tyrosine nitration in proteins by mass spectrometry. J Mass Spectrom 2001;36:616–25.
114. Amoresano A, Chiappetta G, Pucci P, D'Ischia M, Marino G. Bidimensional tandem mass spectrometry for selective identification of nitration sites in proteins. Anal Chem 2007;79:2109–17.
115. Cirulli C, Marino G, Amoresano A. Membrane proteome in *Escherichia coli* probed by MS3 mass spectrometry: a preliminary report. Rapid Commun Mass Spectrom 2007;21:2389–97.

Amino Acids (2011) 40:653–667
DOI 10.1007/s00726-010-0693-1

ORIGINAL ARTICLE

Plasma nitroproteome of kidney disease patients

Marta Piroddi · Angelo Palmese · Francesca Pilolli · Angela Amoresano · Piero Pucci · Claudio Ronco · Francesco Galli

Received: 5 May 2010 / Accepted: 7 July 2010 / Published online: 31 July 2010

Abstract 3′-nitrotyrosine (3NT) is a post-translational modification (PTM) of body fluids and tissues that is sustained by chronic inflammation and oxidative stress, two main clinical traits of chronic kidney disease (CKD). Despite this background, protein targets and their differential susceptibility to in vivo nitration remain almost completely unexplored in CKD. This study reports a first investigation of plasma nitroproteome in these patients, carried out by both immunorecognition and LC-MS/MS techniques. Plasma proteins of chronic and end-stage KD patients showed a higher burden of nitration than in healthy controls, but main nitration targets appeared to be the same in these populations. Immunoblotting data showed that uremic albumin is largely represented in the uremic nitroproteome together with fibrinogen chains (A, B and C), transferrin, α1-antitrypsin, complement factor D, haptoglobin, and IgG light and heavy chains. However, immunopurification and affinity chromatography experiments demonstrated that the relative content of 3NT on the albumin molecule was very low when compared with that of the remaining plasma proteins. The uremic nitroproteome was investigated using also plasma proteins obtained by in vivo ultrafiltration from patients treated with protein leaking or standard high-flux hemodialyzers. The study of these samples revealed the possibility to selectively remove protein nitration products during hemodialysis. Identification of intramolecular sites of nitration was preliminarily obtained in IgG chains isolated by 2D PAGE and assessed by bidimensional tandem mass spectrometry after chemoselective tagging. Further studies are needed to confirm at the molecular level the presence of nitrated Tyr residues in other proteins tentatively identified as nitration targets in this study and to explore the biological meaning of such a selective modification of plasma proteins by reactive nitrogen species in uremia and dialysis patients.

Keywords 3′-nitrotyrosine · Protein nitration · Plasma proteins · Uremia · Hemodialysis · Chronic kidney disease · Uremic toxins · Proteomics · Nitroproteome · Mass spectrometry

Introduction

The covalent product of tyrosine nitration, 3′-Nitro-Tyrosine (3NT), has been measured in the free and protein-bound form in human plasma as biomarker of inflammation and cardiovascular risk (Ptolemy et al. 2007; Parastatidis et al. 2008; Shishehbor et al. 2003; Abello et al. 2009). Available evidence suggests that protein nitration is a pathway and target specific event with consequences on the structure–function relationship of tissue and body fluid proteins (reviewed in Abello et al. 2009; Souza et al. 2008; Ischiropoulos 2009; Peluffo and Radi 2007). Despite this, protein targets and biological consequences of nitration in

Electronic supplementary material The online version of this article (doi:10.1007/s00726-010-0693-1) contains supplementary material, which is available to authorized users.

M. Piroddi · F. Pilolli · F. Galli (✉)
Laboratory of Clinical and Nutritional Biochemistry, Department of Internal Medicine, University of Perugia, Via del Giochetto 06126 Perugia, Italy
e-mail: f.galli@unipg.it

A. Palmese · A. Amoresano · P. Pucci
Department of Organic Chemistry and Biochemistry, University of Naples Federico II, Naples, Italy

C. Ronco
Department of Nephrology, San Bortolo Hospital and International Research Institute of Vicenza, Vicenza, Italy

plasma, i.e. the most investigated biological sample in clinical studies, remain poorly characterized in many conditions. Fibrinogen (Parastatidis et al. 2008; Heffron et al. 2009), ceruloplasmin, transferrin, α1-protease inhibitor, α1-antichymotrypsin, plasminogen (Gole et al. 2000; Pignatelli et al. 2001), and the apolipoproteins Apo B and Apo A1 (Zheng et al. 2004) are in vivo targets, and nitration has been proposed to produce immunogenic effects (Ohmori and Kanayama 2005; Thomson et al. 2007; Parastatidis et al. 2007). These pieces of evidence may suggest a role in coagulation, lipid metabolism, immune and vascular functions, but so far only early evidence of prothrombotic effects by fibrinogen nitration has been provided in literature (Parastatidis et al. 2008). Moreover, only few authors have been able to identify Tyr residues that are modified in vivo by reactive nitrogen species (RNS), which is in turn the conclusive proof of in vivo nitration in a given protein (Abello et al. 2009). In support of a specific and possibly regulated process, albumin as the most abundant serum protein (approx. 60% of total proteins) was not detected as target of nitration in vivo in normal subjects or smokers and lung cancer patients (Pignatelli et al. 2001), as well as in acute respiratory distress syndrome (ARDS) patients (Gole et al. 2000) investigated by immunoblotting methods. Recently, Wayenberg et al. (2009) reported of immunodetectable nitroalbumin in plasma of perinatal asphyxia patients and thanks to a newly developed ELISA they were able to confirm that the relative abundance (mol/mol) of modified and unmodified Tyr residues of albumin is in the order of 10^{-5}. This study pointed out also on the role that kidney damage may have to worsen protein damage by RNS.

This evidence was reported in several other studies that essentially investigated the free form of 3NT as well as other epitopes of protein damage by oxidative and carbonyl stress in uremic plasma (Agalou et al. 2005), while much less attention was paid to characterize protein bound forms. However, signs of protein damage by oxidative/nitrosative damage appear in plasma starting from earliest stages of chronic kidney disease (CKD) (Mitrogianni et al. 2009; Matsuyama et al. 2009) and return to normal after kidney transplantation (Galli 2007; Simmons et al. 2005), thus underlying the key role that the metabolic and depurative function of kidney play on protein metabolism. At the same time, oxidative damage to serum proteins has pathogenic relevance in kidney damage and may thus contribute to the progression of CKD (see Shi et al. 2008 and references therein).

Oxidative stress and protein damage reach their maximal expression in end-stage renal disease (ESRD) and particularly in patients on regular hemodialysis therapy (HD) in which inflammatory and vascular symptoms are also particularly exacerbated (reviewed in Galli 2007; Himmelfarb 2009). In recent years, these aspects have stimulated wide interest on large (or high MW) solutes produced by the abnormal chemistry of ROS/RNS in uremia and dialysis patients, which were increasingly considered for a role as uremic toxins (Vanholder et al. 2006) and for more effective treatment by dialysis therapy (Ward 2005).

Notwithstanding the available information on protein damage by nitration reactions in uremic plasma remains scarce. The extent and targets of nitration during disease progression from pre-dialysis (i.e. in CKD) to dialysis therapy (i.e. in HD patients) as well as the effect of dialysis methods and dialyzer membranes on protein nitration remains unexplored.

As in other clinical studies (recently reviewed by Peluffo and Radi 2007 and by Abello et al. 2009), identification of nitration targets was largely disregarded in uremic plasma. This actually remains an analytical challenge and early studies performed by Mitrogianni et al. (2004) by immunoblotting and immunoprecipitation methods and more recent studies by our group (Galli 2007; Piroddi et al. 2007) carried out by immunoblotting and LC-MS/MS analysis[1] tentatively identified only some of the most abundant targets already described in other inflammatory conditions. However, many other candidate nitration targets are expected to be present in plasma samples of these inflamed patients.

To explore these targets, in this work, we set up a proteomic platform that was applied to the systematic investigation of plasma nitroproteins (the plasma nitroproteome) (1) to preliminarily identifying a broad number of protein targets by 3NT immunodetection, (2) to isolate or enrich these targets for more accurate evaluation and identity assessment by LC-MS/MS, and finally (3) to confirm with selective analysis methods the presence of nitration epitopes within the structure of isolated proteins.

As a first explorative technique in this multi-step protocol, qualitative analysis of nitration targets in uremic and healthy control plasma was carried out by 2D PAGE and WB analyses, which were used also to fractionate immunoidentified nitrated forms to proceed with their identification by LC-MS/MS analysis. The uremic plasma nitroproteome was also investigated on protein samples obtained by in vivo ultrafiltration carried out during HD therapy of patients treated with different dialyzers membranes. Albumin depletion and immunopurification experiments followed by 1D WB analyses were used to selectively enrich 3NT-containing proteins in the uremic plasma. Immunoprecipitation and immunoblotting methods so far validated in previous studies (Gole et al. 2000; Pignatelli et al. 2001; Galli 2007; Mitrogianni et al. 2004) were also applied to further characterize the extent and targets of nitration in uremic plasma.

[1] Piroddi et al. unpublished data, manuscript in preparation. Preliminary data presented at the 11th ICAAP, Vienna 2009.

Unbiased identification of intramolecular targets of nitration was obtained for the first time for a series of selected protein spots corresponding to IgG chain regions on 2D PAGE. This identification was performed by bidimensional tandem MS after chemoselective tagging of nitration epitopes (Amoresano et al. 2007; Chiappetta et al. 2009).

Materials and methods

Patients and blood samples

This observational study has been performed with a cross-sectional design in order to compare levels and targets of protein nitration in pre-dialysis CKD, ESRD patients on hemodialysis therapy (HD), and healthy controls (Ctr). Essential characteristics of these groups are reported in Table 1. Plasma samples of HD patients (5 M/5 F) were obtained from the bank of plasma samples established in our laboratories (samples collected during the years 2007 and 2008 and maintained at −80°C).

Essential criteria to identify CKD patients eligible for the inclusion in the study were a value of creatinine clearance ≤15 ml/min, that is grade IV → V of the KDOQI classification (National Kidney Foundation, Inc 2002; Eckardt et al. 2009). Exclusion criteria were age below to 30 years and higher than 70 years, the absence of clinical and laboratory signs of severe inflammation and protein-energy malnutrition (main laboratory indices considered were plasma levels of IL-6 >5 pg/ml and albumin levels of <3.5 g/dl; these were measured during the first visit for the inclusion in the study), the absence of diagnosis for diabetes, polycystic kidney disease, Lupus and other auto-immune diseases, malignancy and recurrent infections. A final group of 10 CKD patients (6 M/4 F) was identified to meet these criteria and, after providing informed consent, these were included in the study. After inclusion, during one of the morning visits at the ambulatory of Nephrology, 10 ml of heparinized blood were collected from the antecubital vein and plasma was obtained by centrifugation. This was divided in aliquots and stored at −20°C. Blood sampling was performed under overnight fasting conditions.

HD samples were obtained in the middle-week dialysis session from patients on regular three times/week HD. Exclusion criteria were the same than in CKD group plus a dialysis age of lower than 72 months and higher than 9 months. Moreover, only patients with good metabolic control and stabilized on HD therapy with high-flux membranes were considered. Two subgroups of HD patients were thus identified to investigate uremic proteins obtained by in vivo ultrafiltration according with the study design described before in (Galli 2007; Galli et al. 2005). These HD patients were on treatment with protein leaking dialyzers (PLD, $n = 5$) or standard non-protein leaking high-flux polysulfone dialyzers (NPLD, $n = 5$). These two types of high-flux dialyzers differ as regards hollow fiber composition (polymethylmethacrylate vs. polysulfone, respectively) and permeability characteristics (percentage of porosity and pore size result in nominal cut-off values of ≥70 KDa and ≤30 KDa, respectively), which in turn produce different protein patterns in ultrafiltrate samples (Galli 2007). Ultrafiltrate samples were obtained in the first and in the last 5 min of the dialysis treatment by applying a positive pressure on the blood side of the dialyzer to obtain by convection a plasmatic water rich of plasma proteins that was collected from the dialysate chamber by the outlet port in the dialyzer cartridge. Protein concentration in ultrafiltrate samples was checked using the Bradford method and the protein pattern was preliminarily verified by SDS-PAGE and densitometric analysis performed using the software Quantity One (Bio-Rad, CA, USA). Ultrafiltrates used in this study obtained from PLD- and NPLD-treated patients showed total protein levels of 2.4 ± 0.9 and 1.5 ± 0.5 mg/ml, respectively. Further details on the relative abundance and general characteristics of plasma proteins obtained during in vivo ultrafiltration, and dialyzer membrane permeability parameters (percent of porosity and pore size) can be found in (Galli 2007; Piroddi et al. 2007; Galli et al. 2003; Buoncristiani et al. 1999).

Table 1 Clinical characteristics of the CKD and HD subjects included in this study

	Ctr	CKD	HD (PLD)	HD (NPLD)
Age (years)	66 ± 13	72 ± 13	62 ± 12	70 ± 12
Sex (M/F)	5/5	6/4	5/0	3/2
Total proteins (g/dL)	7.0 ± 0.5	6.3 ± 0.5	6.9 ± 0.6	7.0 ± 0.2
Albumin (g/dL)	4.7 ± 0.2	4.6 ± 0.2	4.2 ± 0.3	4.0 ± 0.4
Transferrin (mg/dL)	272.4 ± 48.2	254.6 ± 99.0	291.6 ± 55.9	228.6 ± 157.2
Fibrinogen (mg/dL)	212 ± 26	236 ± 50	248 ± 25	245 ± 42
Creatinineclearance (mg/mL)	89 ± 22	14.9 ± 4.8	–	–
IL-6 (pg/mL)	2.03 ± 0.44	2.47 ± 0.24	5.0 ± 1.9	5.3 ± 2.15
hsCRP (mg/dL)	0.79 ± 3.6	0.77 ± 0.52	2.58 ± 28.6	2.27 ± 33.6

Laboratory determinations

Immunopurification of 3NT-containing proteins and albumin removal by affinity chromatography

Immunoprecipitation (IP) experiments were carried out in whole plasma and ultrafiltrate samples using protein G-Agarose immunoprecipitation magnetic beads (Invitrogen, Milan). The anti-3NT antiserum used was the mouse monoclonal IgG clone 39B6 from Abcam, Cambridge, UK.

Immunoaffinity (IA) of nitroproteins was performed using a Sepharose 4 fast-flow column conjugated with a rabbit polyclonal anti-nitrotyrosine IgG (Affiland, Liège, Belgium) that was used according with the specifications provided by the manufacturer. Nitrotyrosine-containing proteins obtained with this procedure were extensively dialyzed against water and dried down using a speed vacuum concentrator before analysis.

The nitroproteins obtained with IP or IA were resuspended alternatively in SB to run SDS-PAGE and WB experiments or in water to perform LC analyses.

In some experiments, albumin was removed from plasma samples using a blue-agarose gel (Affiland, Liège Belgium) according with the procedure suggested by the manufacturer.

Qualitative analysis of nitroproteins by 1D and 2D PAGE/WB and LC-MS/MS analysis

These analysis methods were used to screen and eventually to isolate potential nitration targets in plasma. Although 3NT immunorecognition has relatively low specificity, it remains an essential step of nitroprotein analysis since it allows to restrict the number of candidate targets in complex protein samples to be further investigated by higher levels of accuracy and precision for protein identity and nitration epitope confirmation (see chemoselective labeling and LC-MS/MS analysis protocols described below). For 1D WB analysis, whole plasma or ultrafiltrate proteins suspended in SB were incubated for 5 min at 100°C, and 15 µg of proteins were loaded in each lane of a monodimensional polyacrylamide gel (1D-PAGE; 10% of T) that was run using a mini Protean Cell System (Bio-Rad). After 1D-PAGE, proteins were transferred onto a nitrocellulose membrane using a trans-blot SD semi-dry transfer cell (Bio-Rad), and were incubated with the same mouse monoclonal IgG used for IP (clone 39B6 from Abcam, Cambridge, UK). Thus, a further incubation step was performed using a secondary goat anti-mouse peroxidase-conjugated antibody (Millipore, CA, USA). In some experiments nitroproteins were also immunodetected using a rabbit polyclonal anti-3NTprimary antibody (Sigma-Aldrich, Milan, Italy) and a goat anti-rabbit peroxidase-conjugated antibody (Millipore, CA, USA). The membrane was washed three times (5 min each) in TBS containing 0.1% vol/vol Tween 20 and the signal was developed on BioMax light films (KodaK, NY, USA) using an ECL solution containing 1 M Tris–HCl pH 8.5, 1.25 mM luminol in DMSO, 0.25 mM cumaric acid in DMSO, 0.015% vol/vol H_2O_2. Image analysis and densitometry were performed using Gel-Pro Analyzer 3.1 (MediaCybernetics, Inc., MD, USA).

Specificity of anti-3NT monoclonal antisera used in this study was tested by protein reduction with $Na_2S_2O_4$ according with the procedure previously reported in (Chiappetta et al. 2009). Other tests included the saturation of the incubation medium containing the first antibody with free 3NT and the omission of the primary antibody. These tests were performed with either low or high loads of plasma proteins during SDS-PAGE (7.5 or 15 µg/lane, respectively) and with different times of exposure of light films to ECL (from 10 to 60 s.).

For 2 PAGE and WB experiments, proteins were precipitated overnight in ice-cold acetone and the precipitate was centrifuged for 2 min at 14,000×*g*, 4°C. The final pellet (300 µg of proteins) was dissolved in rehydration buffer and then was diluted to the appropriate final protein concentration in 350 µl of DTT 60 mM, Tris 40 mM, urea 7 M, thiourea 2 mM, CHAPS 4%, Bio-Lyte ampholite (Bio-Rad) 0.7 µl, and bromophenol blue. The samples were loaded onto immobilized pH gradient gel strips (17 cm, pH 3–10 NL), and isoelectric focusing was performed using a Protean IEF cell system (Bio-Rad Labs). The second dimension was performed on 9 to 16% polyacrylamide linear gradient electrophoresis gels using a Protean II xi Cell (Bio-Rad). 2D WB was performed using the same procedure described above for 1D WB using a Trans-blot cell (Bio-Rad). 2D WB and PAGE image matching was performed by PD Quest software (Bio-Rad), selected spots were picked and in-gel trypsinization was carried out as described in (Sultana et al. 2006). Tryptic digest was evaluated by nanospray LC/MSMS using a an HCT nano-HPLC-Chip/MS system Agilent 1200 Series equipped with a tridimensional ion trap system (Agilent Technologies, Milan, Italy). A customized chip program for shotgun proteomics was applied according with the manufacturer. 2D map matching for protein identification was performed using the resources available in the SWISS-PROT database (http://www.expasy.org/cgi-bin/map2/def?PLASMA_HUMAN).

Semiquantitative analysis of protein 3NT by immuno-dot blot analysis (DB)

Plasma samples were prepared for the analysis by appropriate dilution in Laemli buffer (Tris–HCl 0.12 M pH 6.8,

glycerol 21.7%, SDS 4.28%), and different amounts of total proteins were spotted onto a nitrocellulose sheet that was soaked in Tris–HCl 20 mM pH 7.5, NaCl 500 mM, and then mounted for the analysis onto a Bio-Dot microfiltration apparatus (Bio-rad Laboratories s.r.l., Milan). Immunoblot and image developing were obtained with the same procedure described above for 1D WB experiments.

The assay was calibrated using NO-BSA that was prepared according with (Amoresano et al. 2007). Briefly, BSA (10 mg/mL) in 200 mM Tris buffer (pH 8.0) was exposed to tetra-nitro-methane (Sigma-Aldrich) at the final concentration of 350 mM. The reaction mixture was stirred at room temperature for 30 min, and NO-BSA was rapidly desalted by size exclusion chromatography on a Sephadex G-25 M column (Sigma-Aldrich). Protein elution was monitored at 280 and 360 nm, and the protein-containing fraction was collected and stored at −20°C. The content of 3NT of this nitrated protein was checked by MALDI and nano-LC-MS/MS analysis according with the method reported in (Chiappetta et al. 2009). Assuming from MALDI data that approximately 30% of albumin exposed to TNM is nitrated, and that 4 Tyr residues per molecule can be converted to 3NT (Chiappetta et al. 2009), a value of 20 pmol of 3NT/μg of NO-BSA was considered to make calibration curves. Incremental concentrations of this reference protein were loaded in each DB experiment (in the range 0.4–160 pmol/spot that correspond to 0.02–8 μg/spot) and the assay response with these protein loads was verified by linear regression analysis within each analysis session. Linearity was also tested assessing three different times of exposure to ECL. With NO-BSA linearity ($R = 0.988$, $p < 0.01$) was achieved in the range 0.4–16 pmol of protein/spot with time of exposure to ECL of 15 s. Under these conditions, the observed detection limit was 0.2 pmol NO-BSA/spot. To further increase analysis precision, in the same session each DB experiment was run in duplicate. A first DB run was performed to calibrate protein load and ECL exposure (usually 0.5 min) and a second DB run was performed to confirm optimal conditions and to obtain final semi-quantitative data. In each of these two DB experiments, three incremental concentrations of plasma proteins were spotted in series. The amount of 3NT in plasma proteins was expressed in pmol/μg of proteins. An intra-assay CV % of ≤10% was observed for NO-BSA and of ≤19% for plasma proteins.

Reduction of nitrated plasma proteins to secondary amines by $Na_2S_2O_4$ (dithionite or sodium hydrosulfite) (Sigma-Aldrich) was used to confirm specificity of anti-3NT IgG used in these DB experiments by a stripping–reprobing procedure. Briefly, proteins transferred onto blotting membranes were stripped using the Restore™ stripping buffer (Thermo Scientific, IL, USA) and reduced incubating the membrane at room temperature in 20 mM dithionite dissolved in 50 mM borate pH 9.5 for 5 min. Reprobing was thus performed as described above using the same primary anti-3NT antibody.

Gel-free bidimensional tandem mass spectrometry analysis of nitroproteins after chemoselective tagging

Unbiased semiquantitative analysis of nitroproteins in human plasma was performed by chemoselective tagging and bidimensional nano-LC-MS/MS according with the procedure previously described in (Amoresano et al. 2007). Briefly, after 2D PAGE of plasma proteins, selected spots in the region corresponding to IgG lambda and heavy chains were isolated and processed for trypsinization. Reduction of nitro groups to amino groups in the trypsin digest was performed by $Na_2S_2O_4$ at a molar ratio of 100:1 with expected 3NT content as calculated in NO-BSA (Amoresano et al. 2007; Chiappetta et al. 2009). Then, 3-aminotyrosines were labeled with a solution of dansyl-chloride (DNS-Cl) 18.5 ng/μL in ACN (1,000-fold molar excess) and the peptide mixture was analyzed by LC-MS/MS analysis using a 4000Q-Trap machine (Applied Biosystems) coupled to an 1100 nano-HPLC system (Agilent Technologies). Total ion counts (TIC), precursor ion scan (PIS) and MS3 analyses were performed and resulting data were acquired using Analyst software (Applied Biosystems) and further processed by MASCOT in house software as described in (Amoresano et al. 2007).

Statistics

Results were expressed as the mean ± SD. Statistical analysis and data plotting were carried out using Microcal™ Origin 7.5 (Microcal Software Inc., Northampton, MA, USA) and Sigma Plot 9.0 (Systat Software Inc., San Jose, CA, USA). Probability values of <0.05 were considered as significant.

Results

Semi-quantitative analysis of nitroproteins by dot blot

Figure 1 shows a typical semiquantitative DB analysis carried out to measure nitroproteins in plasma of CKD, HD and Ctr groups. All the DB experiments were optimized as regards optimum of protein load, antibody performance and ECL exposure. The resulting photographic images (Panel A, left) were digitalized and analyzed by densitometry against a calibration curve of nitro-BSA (Panel B). Each DB experiment was tested also for specificity by reduction with dithionite (Panel A, right). Panel C shows semi-quantitative data obtained with this test. After correction of unspecific

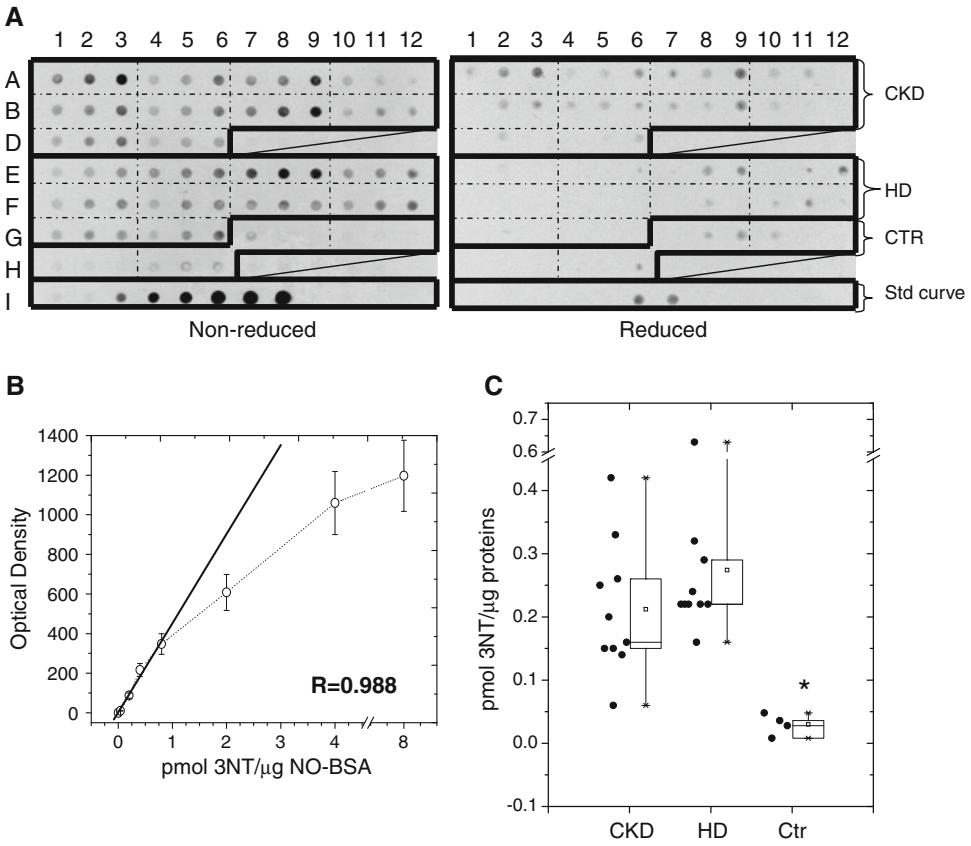


Fig. 1 Immuno-dot blot analysis (DB) of 3′-nitro-Tyr (3NT)-containing proteins in chronic kidney disease (CKD), hemodialysis (HD) and healthy control (Ctr) plasma. This test was carried out using an improved version of the procedure described by Pignatelli et al. (2001). 3NT immunodetection was performed with the method described in detail in the section "Materials and methods". DB tests (**a**) were performed by loading for each plasma sample three incremental amounts of total proteins (from 5 to 20 µg) that were selected in order to achieve a linear response during film exposure to ECL. Specificity in the immunorecognition of 3NT was checked by reduction with sodium dithionite (**a**, right). After densitometry, 3NT concentrations were calculated using a calibration curve of nitro-bovine serum albumin (NO-BSA; **b**). The linear portion of the calibration curve shown by the *straight fitting line* was used to calibrate the assay. Mean levels of 3NT in CKD, HD and Ctr groups are shown in **c** (as pmol 3NT/µg of total proteins). $*p < 0.01$ versus CKD and HD

recognition revealed by dithionite reduction (this was particularly marked in CKD group), the levels of immunodetectable 3NT in CKD and HD plasma proteins were determined as 0.212 ± 0.104 and 0.274 ± 0.132 pmol 3NT/µg total proteins, respectively (p = NS). A high inter-individual variability was observed in these two groups of patients (Panel C). In Ctr, a four-time higher protein load per spot than in uremic patients was needed to measure ECL signals that linearly corresponded to the amount of proteins assessed in the DB test (not shown), and a value of 0.030 ± 0.017 pmol 3NT/µg total proteins was calculated ($p < 0.05$ vs. both CKD and HD groups).

Qualitative analysis of nitrated proteins

Plasma samples from uremic patient and healthy controls were first analyzed by 1D PAGE and WB experiments carried out under saturating conditions (obtained with a load of 10 µg of total proteins per lane) to explore potential 3NT-containing proteins (shown in the supplemental material as Fig. 1a) that will be further investigated to confirm protein identity and nitration sites by LC-MS/MS (see below). The specificity of the monoclonal antisera used in this study was confirmed by applying different analytical conditions as shown in Fig. 1b (see the supplemental files).

Although this analysis provides very limited resolution, abundant immunodetectable material was recognized in both the uremic (only HD patients were assessed) and healthy control plasma. Main nitrated species could be identified in the region between 50 and 80 KDa, while another series of bands with lower intensity were present in the 100–260 KDa region. Finally, an intense band was observed just below 30 KDa. The conditions used to obtain this qualitative image (protein load and exposure time to ECL were maximized) interfered with the possibility to obtain semiquantitative information at least for most abundant nitrated species.

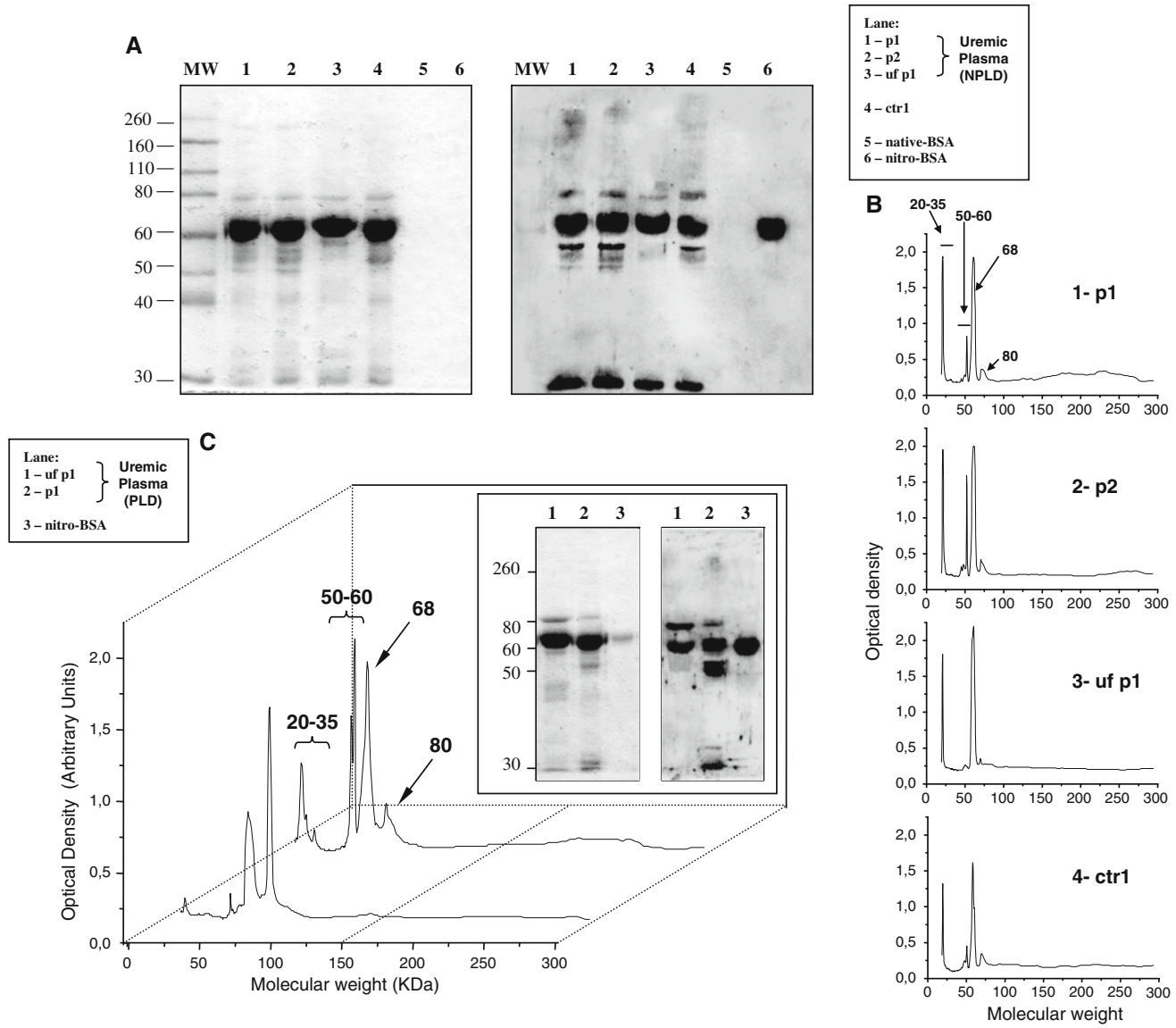


Fig. 2 Qualitative analysis of 3NT-containing proteins in HD plasma and ultrafiltrate samples. 1D-WB analyses were carried out in PLD and NPLD plasma proteins under non-saturating conditions (7.5 μg of proteins/lane, ECL exposure 0.5 min.). **a** The SDS-PAGE and WB analysis carried out in NPLD plasma and ultrafiltrate samples, representative of the entire set of samples investigated in this study. Corresponding densitometric data are shown in **b** and **c**. Densitometry profiles of WB analyses carried out on PLD plasma and ultrafiltrate are shown in detail (**c**) to highlight the enrichment of the protein band with apparent MW = 80 KDa. NPLD ultrafiltrate (**b**, Panel 3) shows a marked increase of the band in the region between 25 and 30 KDa. Abbreviations are as in Fig. 1

Figure 2a shows the 1D WB analysis carried out on whole plasma and ultrafiltrate samples from patients treated with PLD and NPLD dialyzers. Analyses were performed under the same conditions described in Fig. 1, but with lower protein loads (7.5 μg of total proteins per lane) to avoid signal saturation and thus to tentatively increase resolution in the region between 50 and 80 KDa, i.e. the region in which the most abundant nitrated proteins were observed.

Densitometry scanning of 1D WB images showed very similar patterns of nitrated proteins in uremic and healthy control plasma (Fig. 2b, panels 1, 2 and 4). Although semiquantitative analysis of immunorecognized species by this test has limited reliability, all the assessed samples of uremic plasma showed a higher area under the curve than in the healthy controls, thus confirming DB results shown in the previous paragraph.

On the basis of apparent MW, qualitative results obtained by 1D WB experiments were consistent with the tentative identification of albumin as main target of nitration in uremic and healthy control plasma (Fig. 2b, peak at 68 KDa). The peaks with apparent MW between 50 and 60 KDa were tentatively assigned to the already reported nitration targets fibrinogen chains, α1-antitrypsin and α1-antichymotrypsin (Parastatidis et al. 2008; Heffron et al.

2009; Gole et al. 2000; Pignatelli et al. 2001). Finally, the nitrated peak with MW of 80 KDa could be compatible with transferrin, while the peaks between 20 and 35 KDa were consistent with the reported nitration of Apo-A1 (Zheng et al. 2004), but also with a series of other proteins including the IgG light or λ chain, one of the most abundant protein fractions in plasma.

When examined by 1D WB, ultrafiltrate samples from patients treated with NPLD (Fig. 2b, panel 3) and PLD (Fig. 2c) showed the presence of a less complex pattern of nitrated proteins with respect to whole plasma. In particular, NPLD ultrafiltrates essentially showed the presence of peaks at 68 KDa and in the region 20–35 KDa, while PLD ultrafiltrate contained peaks with MW of 68 and 80 KDa. In both ultrafiltrate samples, the bands corresponding to nitrated proteins occurring in the 50–60 KDa region were markedly reduced with respect to whole plasma.

2D PAGE and LC-MS/MS identification of nitrated proteins

Figure 3 shows typical 2D PAGE analyses and the corresponding WB images obtained with anti-3NT antisera of uremic proteins in plasma (Fig. 3a), and in PLD and NPLD ultrafiltrate samples (Fig. 3b) that were previously examined by 1D WB (see above and Fig. 2). After matching WB and coomassie stained 2D PAGE patterns (Fig. 3a), immunorecognized nitrated proteins were isolated by cutting gel spots, and these candidate nitroproteins were submitted to molecular identification by in situ digested with trypsin and LC-MS/MS analysis of the resulting peptide digests. Spectral data were used to search for a non-redundant protein database using the Mascot software and identified proteins are shown in Table 2. Most abundant plasma proteins such as albumin, IgG light and heavy chains, antithrombin and α1-antitrypsin were tentatively identified with this procedure as nitration targets. However, also some low abundant proteins like the complement factor D could be identified. Other candidate targets of nitration were assigned by direct matching of WB images with existing 2D PAGE images of human plasma available for consultation and analysis at: http://www.expasy.org/cgi-bin/map2/def?PLASMA_HUMAN). These proteins included fibrinogen A, B and C chains that showed their characteristics sequel of spots in the 2D MW/PI region, serotransferrin, ceruloplasmin and haptoglobin (Fig. 3a; Table 2). The same analyses carried out on selected spots from 2D-PAGE gels of PLD and NPLD ultrafiltrate samples (Fig. 3b), produced essentially comparable results.

This LC-MS/MS identification of nitrated proteins fractionated by 2D-PAGE was useful to further increase the level of information obtained in the preliminary 1D WB experiments performed on whole plasma and ultrafiltrate samples shown in Fig. 2. Thus, the peak at 68 KDa (Fig. 2b, c) very likely corresponds to albumin and the peak region of 20–35 KDa should include IgG light chains together with other minor components as complement factor D. In the same way, the peak at 80 KDa essentially consisted of transferrin, which is highly enriched in PLD ultrafiltrate. The findings described for the 1D WB analysis of ultrafiltrate samples were also confirmed by this analysis. Main nitrated forms of the region 50–60 KDa are absent in these samples (essentially fibrinogen A and B chains, and IgG heavy chain) while α1-antitrypsin was detected together with other minor components in the PI region of albumin that was not identified.

Immunopurification and albumin depletion

In an attempt to selectively enrich the relative concentration of plasma nitroproteins, immunoaffinity (IA) and immunoprecipitation (IP) experiments (Fig. 4a, b, respectively) were performed using the same antisera used in the WB and DB experiments of above. The IA procedure produced similar protein recoveries in uremic and healthy control subjects suggesting that the experiment was performed under saturating conditions with respect to the binding capability of anti-3NT antisera. Accordingly, the fractions not retained during the immunopurification process contained abundant anti-3NT positive material (Fig. 4c). The IP experiment was then carried out under non-saturating conditions providing a higher purification yield than IA (Fig. 4b). Indeed, nitroproteins were not detectable in the washing buffer and in the supernatants after IP (Fig. 4c), and the amount of 3NT-containing proteins was lower in healthy controls than in HD patients (mean difference $= 22\%$; $n = 4$ each).

IA and IP samples prepared from both whole plasma and ultrafiltrate samples were then analyzed by 1D PAGE and WB showing that most abundant nitrated components have an apparent MW between 50 and 60 KDa. This finding confirms the elective nitration of proteins such as fibrinogen chains that were described in the 2D WB analysis of Fig. 3 and in literature (see Parastatidis et al. 2008; other references are reported in the "Discussion" section). In this MW region, nitration of α1-antitrypsin and IgG heavy chains had also been tentatively inferred by 2D WB (Fig. 3). At the same time, albumin does not appear to be recovered during IA and IP experiments. This was particularly obvious in the IA experiment of Fig. 4a. This finding was confirmed by albumin depletion of plasma samples carried out with blu-agarose affinity chromatography and subsequent SDS-PAGE and DB analyses (shown in the supplemental files as Fig. 2a). In both uremic or healthy control plasma, the large majority of the immunodetectable 3NT (approx. 95% of the signal detected following dithionite reduction) was observed in albumin-depleted

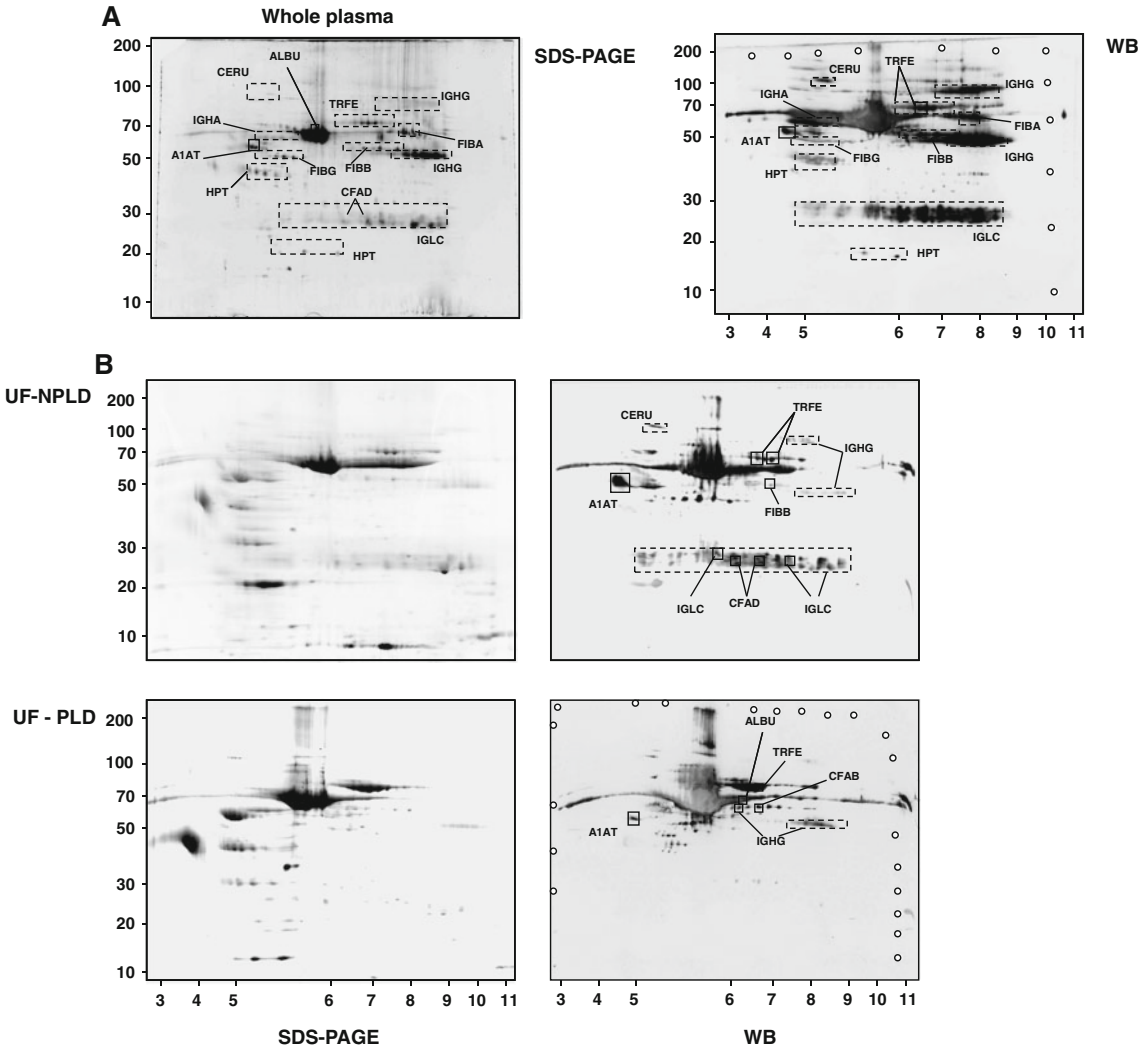


Fig. 3 2D-PAGE and WB analysis for the qualitative evaluation of 3NT-containing proteins in HD plasma and ultrafiltrate samples. Qualitative identification of nitrated proteins in HD plasma and ultrafiltrate samples was carried out by 2D-PAGE (**a** top panel, **b** left panels) and WB (**a** bottom panel, **b** right panels) according to the procedure previously described in (Galli 2007; Piroddi et al. 2007). Protein identification by LC-MS/MS analysis was carried out in selected protein spots (highlighted by *squared straight line*). Other proteins were identified by gel matching (spots highlighted by *squared dashed line*) using as template the human plasma 2D map available for consultation at the Expasy website (http://www.expasy.org/cgi-bin/map2/def?PLASMA_HUMAN). The list of proteins identified with this procedure is reported in Table 2. *White dots* shown in WB films indicate reference points used during image matching with the corresponding 2D PAGE by PD Quest software (Bio-Rad)

samples. As calculated by densitometric analysis of SDS-PAGE, these samples contained approximately 45% of the total proteins (shown as supplemental Fig. 2b). These data demonstrate that albumin is much less extensively nitrated than fibrinogen and other less abundant species in plasma.

Bidimensional MS analysis of nitroproteins after chemoselective labeling and identification of 3NT residues

Direct identification of nitrated proteins and accurate definition of nitration sites in uremic plasma proteins was inferred by bidimensional tandem mass spectrometry (Amoresano et al. 2007) after trypsinization and dansyl-chloride labeling of nitropeptides. Selected spots of 2D PAGE maps and immunoresponsive to anti-3NT antisera during 2D WB experiments were assessed with this procedure. After selective labeling of 3NT residues by reduction to 3-aminotyrosine and specific dansylation, the exact location of original nitro-groups within the amino acid sequence of nitropeptides was successfully identified by bidimensional tandem MS analysis in some components of the IgG chain region (Table 3). This MS analysis was performed by PIS and MS3 scan mode combination.

Table 2 Tentative identification of nitroprotein targets in uremic plasma as inferred from 2D maps and WB analyses

Spot Nr	Protein	SP code	MW	Nr peptides
LC-MS/MS analysis				
A1AT	Alpha 1-antitrypsin	P01009	46878	37
ANT3	Antithrombin-III	P01008	53025	4
ALBU	Serum Albumin	P02768	71317	39
CFAD	Complement Factor D	P00746	27529	3
IGKC	Ig K chain C region	P01834	11773	9
–	Ig K chain V–III region WOL	P01623	11853	3
Gel-matching procedure				
IGHG	Ig gamma chain C region	P01860		
IGLC	Ig lambda chain C region	P01842		
FIBA	Fibrinogen alpha chain	P02671		
FIBB	Fibrinogen beta chain	P02675		
FIBG	Fibrinogen gamma chain	P02679		
CERU	Ceruloplasmin	P00450		
HPT	Haptoglobin	P00738		
TRFE	Serotransferrin	P02787		
IGHA	Ig alpha-1 chain C region	P01876		

In PIS mode, the MS detector can only detect peptide precursor ions producing the *m/z* 170 fragment. The selected precursor ions were then subjected to a combined MS2 and MS3 experiment to specifically detect only those ions originating the dansyl-specific *m/z* transition 234 → 170 in MS3 mode. Figure 5 shows a typical profile obtained from this analysis carried out in some of the main spots corresponding to the IgG chain regions in the 2D map shown in Fig. 3a. As expected, the TIC chromatogram (Fig. 5, upper panel) exhibited a larger number of signals most of which were very likely unrelated to 3NT-containing peptides. In fact, the corresponding precursor ion TIC profile shown in the lower panel, exhibited a much simpler ion pattern in which the MS2 spectra showed only the diagnostic fragment ion at *m/z* 170. Further selection based on MS3 scan recording the *m/z* transition 234 → 170 (Amoresano et al. 2007) removed the remaining false-positives leading to the identification of truly nitrated peptides. Finally, MS/MS analysis of the labeled peptides produced fragmentation spectra that were used to determine peptide sequences for unbiased evaluation of protein identity and unambiguous location of the nitration site within its structure. The results of this PIS and MS3 procedure are reported in Table 3 that shows modified Tyr residues in chemoselective-labeled nitropeptide and the SwissProt accession code for protein identification of seven different peptides all belonging to immunoglobulin species.

Discussion

NO-mediated protein damage may contribute to sustain events of uremic toxicity and the cardio-inflammatory comorbidity of CKD patients. Actually, protein nitration is increased in uremia as well as in other inflammatory conditions and may produce functional consequences that are compatible with an increased cardiovascular risk (Ischiropoulos 2009). Notwithstanding, protein nitration in uremic plasma remains poorly characterized as levels, targets and clinical importance are concerned. The main reason for this lack of information is the complexity and poor reliability of methods and techniques currently available to perform these studies. Extensive characterization of 3NT-containing proteins in complex biological samples, such as plasma, remains in fact an analytical challenge and even most updated and sophisticated shotgun proteomics techniques fail to provide suitable (high-throughput and selective) analytical platforms (Abello et al. 2009). The proteomic platform developed and applied in this study to investigate qualitative and semi-quantitative aspects of uremic plasma nitroproteome includes low-specificity exploration tools such as 2D-PAGE, immuno-detection/isolation techniques, and also specific tools for the unbiased identification of 3NT-containing species by bidimensional MS after chemoselective labeling.

The immuno-dot blot analysis method originally proposed by Pignatelli et al. (2001) was improved and used to obtain a semi-quantitative assessment of total nitroproteins in plasma before to proceed with qualitative assessments. Although this analysis method overestimates the actual levels of 3NT in plasma proteins (Ryberg and Caidahl 2007), it has been useful to demonstrate an increased burden of nitration in uremic plasma than in healthy control plasma. Indeed, CKD and HD plasma contained 0.212 ± 0.104 and 0.274 ± 0.132 pmol 3NT/μg total plasma proteins, respectively, while healthy control plasma contained 0.030 ± 0.017 pmol 3NT/μg total proteins ($p < 0.01$ vs. both the two groups of patients).

In the exploration of uremic nitroproteins (qualitative analysis) carried out by 1D and 2D WB, albumin appeared the most abundant plasma nitroprotein; however, immunopurification experiments and selective removal of this protein showed that the relative contribution of the nitrated form of albumin to total protein 3NT levels is very low. These findings are consistent with the results obtained in earlier studies carried out by some of us on uremic proteins obtained by in vivo ultrafiltration (Galli 2007; Piroddi et al. 2007) and also with the results obtained by other Authors that studied whole plasma and also anti-3NT immunoprecipitated uremic proteins (Mitrogianni et al. 2004).

Albumin was not detected as nitration target by WB even after immunoprecipitation of 3NT-containing plasma

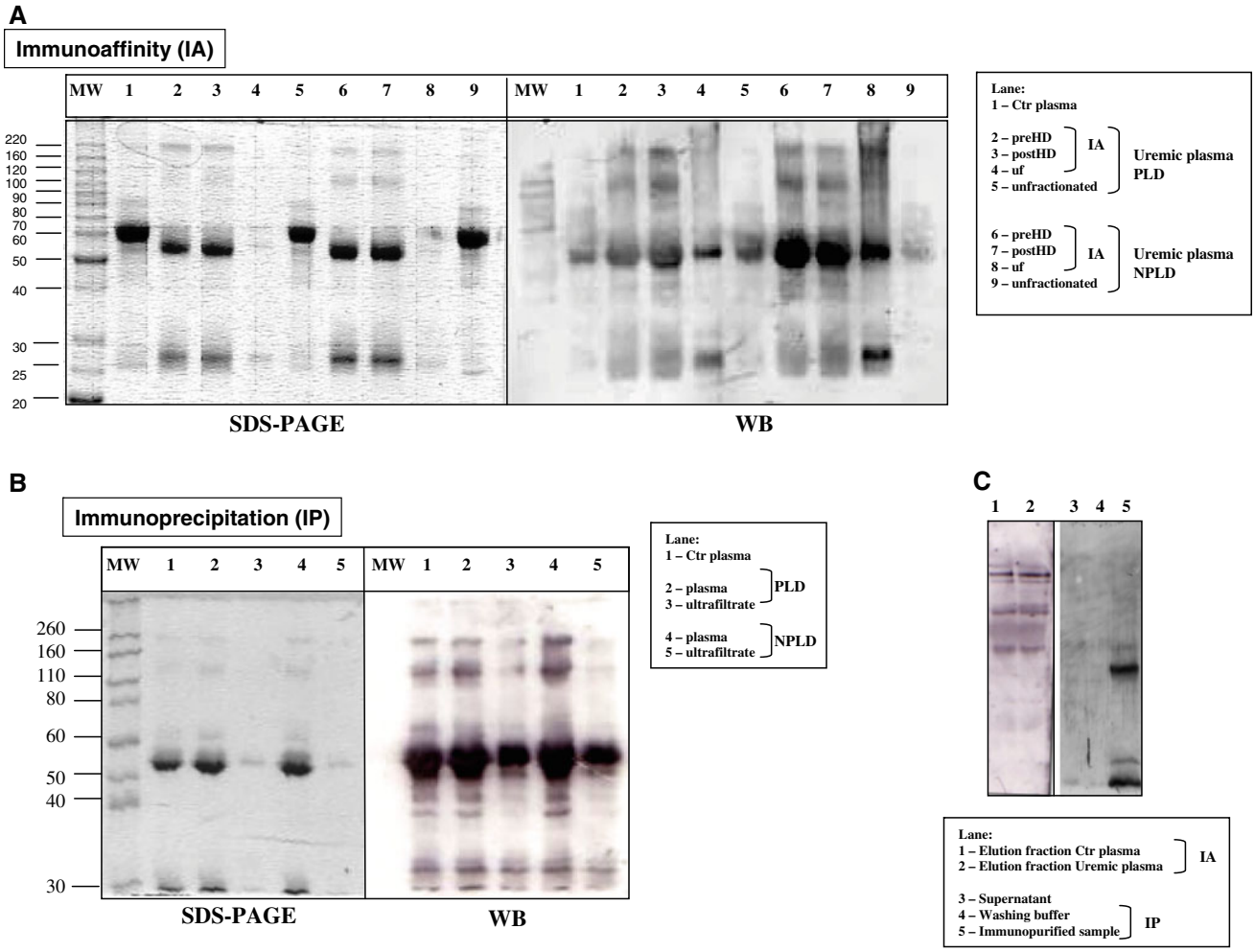


Fig. 4 Partial purification of nitrated species in uremic plasma by immunoaffinity and immunoprecipitation. Pooled samples of HD and healthy control plasma were treated for IA or IP analysis as described in the text and the resulting protein samples were assessed by 1D PAGE (15 μg proteins/lane) and WB analysis (**a** and **b**, respectively). WB analysis of the nitrated proteins recovered in the elution fractions of IA experiments and the supernatant and washing buffer obtained from IP experiments are shown in **c**. Further details are reported in the text. Abbreviations are as in Fig. 1

proteins in the other observational studies of Gole et al. [2000] and Pignatelli et al. [2001]. Accordingly, no identification of intramolecular targets was reported for the in vivo nitrated form of albumin that after in vitro exposure to peroxynitrite was identified to contain modified Tyr residues in position 138 and 411 (Jiao et al. 2001). On the contrary, nitration sites have been identified in the in vivo nitrated form of fibrinogen β chain (Parastatidis et al. 2008), which is proved to be one of the most nitrated species in human plasma in spite of a relative abundance on total proteins of 3%. Thus, albumin appears to sustain the unspecific immunorecognition of 3NT in plasma proteins, and this may explain also why in the absence of nitroprotein separation, immunodetection methods largely overestimate total nitroprotein levels in plasma.

Immunodetectable nitroalbumin has been recently reported and quantitated with a newly developed ELISA test, in plasma of neonates with asphyxia-induced encephalopathy by Wayenberg et al. (2009). These authors observed in patients with neurological symptoms median levels of nitroalbumin of 14.4 ng/ml that decreased to 7.3 ng/ml in asymptomatic neonates, which correspond to approximately 3.7 and 1.8×10^{-5} mol nitroalbumin/mol human serum albumin (HSA), respectively. These values increased to approximately 5.7×10^{-5} mol nitroalbumin/mol HSA (i.e. 22.6 ng/ml) in patients that in combination with neurological symptoms developed severe renal injury, but not hepatic or cardiac complications. These results confirm the key role that the kidney plays to control plasma protein damage and also the substoichiometric character of albumin nitration that in turn reflects the low levels of protein 3NT and the selectivity of nitration pathways in human plasma (Abello et al. 2009; Souza et al. 2008; Ischiropoulos 2009; Ryberg and Caidahl 2007).

Table 3 Bidimensional tandem mass spectrometry identification of nitration sites in protein spots corresponding to main IgG regions of 2D PAGE (Fig. 3)

Protein	SP code	Nitration site
Ig lambda chain V–I region BL2	P06316	VTISCSGSSSNIGNDY*V
Ig lambda chain V–I region EPS	P06888	VSISCSGSSSNIGKNYVDWY*QQLPGTAPK
Ig heavy chain V–II region ARH-77	P06331	YYY*GMDVWG
Ig heavy chain V–III region BUR	P01773	TYY*ADSVRGR
Ig heavy chain V–III region VH26	P01764	DNSKNTLY*LQMNSLR
		STYY*ADSVKGRFTISRDNSK
		GLEWVSAISGSGGSTYY*ADSVK

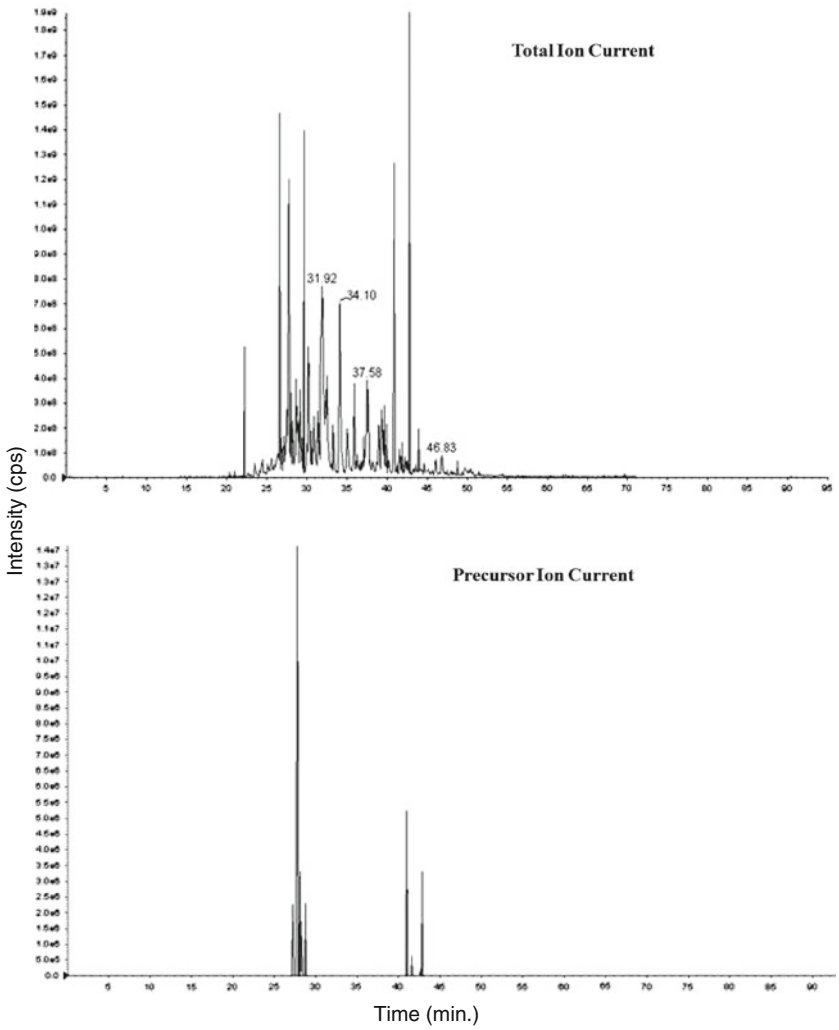


Fig. 5 Bidimensional LC-MS/MS analysis of IgG nitropeptides after chemoselective labeling. Tryptic peptide mixtures from selected spots occurring in the IgG regions of the 2D map and immunoresponsive to anti-3NT IgG were reduced and the resulting aminotyrosine epitopes of nitropeptides were labeled by dansyl-chloride and processed for 2D-LC-MS/MS analysis according to (Amoresano et al. 2007). The *upper* and *lower* panels show the total ion current (TIC) chromatogram and the precursor ion scan monitoring of the dansyl-chloride fragment ion with *m/z* 170, respectively. These ion chromatograms are representative of the entire set of 2D-LC-MS/MS analyses performed to identify nitropeptides of IgG chains reported in Table 3

During preliminary immunorecognition tests (qualitative analysis by 2D PAGE and WB, Fig. 3) other nitration targets identified in uremic plasma corresponded to protein targets previously described in small observational trials on other inflamed patients, as ARDS and lung cancer patients (Gole et al. 2000; Pignatelli et al. 2001). These included transferrin, ceruloplasmin, fibrinogen, and α1-antitrypsin. At the same time, other targets were tentatively identified for the first time in this study by 2D PAGE/WB and LC-MS/MS analysis and included IgG light and heavy chains, complement factor D, and haptoglobin.

Fibrinogen β chain, the most characterized nitration target in human plasma (Parastatidis et al. 2008; Nowak et al. 2007), was confirmed as one of the main nitrated species also in uremic patients. However, this nitrated form of fibrinogen was present also in healthy control plasma and the present study suggests also that α and γ chains of fibrinogen are nitrated in uremic and healthy control

plasma. Thus, fibrinogen chains seem to represent constitutive targets of nitration in human plasma and future studies are required to quantify their extent of nitration in patients and healthy controls, since this could reveal a higher risk of thrombogenicity (Parastatidis et al. 2008; Peluffo and Radi 2007; Heffron et al. 2009; Nowak et al. 2007). Fibrinogen nitration was recently proposed to represent a surrogate marker of inflammation of possible relevance in clinical trials (Heffron et al. 2009) and thus could be adopted in laboratory protocols to assess sub-clinical inflammation in uremia and dialysis patients. Further studies should be addressed to establish whether nitration of coagulation proteins is part of the oxidative stress-mediated events that are believed to sustain a defective coagulation (increased fibrinolytic activity) of uremia and dialysis patients (Pawlak et al. 2006). Thus, the study of fibrinogen nitration and more in general of other nitration targets among coagulation proteins (these include for instance α1-antitrypsin, which is abundantly expressed in the nitroproteome of uremic plasma and NPLD ultrafiltrate), could be of relevance in uremia and dialysis patients since chronic inflammation and impaired coagulation are main uremic symptoms that cannot be corrected with present dialysis therapies.

The study of the uremic nitroproteome performed on plasma proteins separated by in vivo ultrafiltration revealed first information as regards the potential of extracorporeal treatments to correct the accumulation of large solutes containing RNS-derived post-translational modifications (PTMs). These particular plasma protein samples were obtained from patients on treatment with PLD and NPLD dialyzer membranes. These two types of high-flux dialyzers produce different protein patterns in ultrafiltrate samples due to percentage of porosity and pore size characteristics that result in nominal cut-off values of $\geq$70 KDa and $\leq$30 KDa, respectively (Galli 2007). In this study, we found that albumin is the most abundant component recognized by 3NT immunoblotting in both the PLD and NPLD ultrafiltrate samples, followed by IgG and fibrinogen chains, transferrin, α1-antitrypsin, ceruloplasmin and haptoglobin. Transferrin is one of the most abundant nitrated forms selectively removed by PLD. This is the most important pool of iron in the body with the highest turnover rate for this transition metal, and noteworthy dialysis patients show a characteristic tendency to develop disturbances of transferrinemia and impaired transferrin saturation (Canavese et al. 2004). Another target of nitration in uremic plasma associated with iron metabolism is ceruloplasmin, which was observed in ultrafiltrate samples from NPLD-treated patients, while only traces of this protein were observed in PLD ultrafiltrates. The nitration of this protein was already identified by immunoprecipitation and WB analysis of HD plasma in the pioneering study of Mitrogianni et al. (2004). These authors identified only this target as more extensively nitrated in HD patients than in healthy controls. Nitration of these two iron related proteins in uremic plasma might have implications for an abnormal metabolism of this transition metal, for uremic anemia and for a higher susceptibility to oxidative stress (Sezer et al. 2007). A defective iron metabolism increases also the susceptibility to infections and is a well recognized cause of impaired immune function in CKD and HD patients. These clinical traits are often associated with iron overload by massive iron therapy that is used to combat severe anemia in HD patients (Canavese et al. 2004).

NPLD were found to selectively remove α1-antitrypsin and IgG light chains, which are heavily represented in the nitroproteome of uremic plasma. On the contrary, IgG heavy chains, another main nitration target in uremic plasma, are not present in ultrafiltrate samples. Key components of complement pathways as the B and D factors have been observed in traces in the nitroproteome of NPLD ultrafiltrates. Complement factor B and D are part of the alternate pathway of the complement system that, together with complement factor 3b, generate the C3 or C5 convertase. Complement activation during HD reflects the scarce biocompatibility of the extracorporeal treatment and may have severe clinical consequences for these patients (Koller et al. 2004). Changes in IgG and complement factors in uremic plasma induced by HD therapy have been proposed to sustain leucopenia and neutrophil activation with mechanisms that involve oxidative stress (Koller et al. 2004; Kormoczi et al. 2001). Nitration of both IgG and complement factors has been investigated in vitro. IgG nitration results in a defective binding of C1q subcomponent of human complement (McCall and Easterbrook-Smith 1989), while nitration of C5a results in a lowered biological activity by reduced binding to granulocyte receptor (Johnson and Chenoweth 1985).

Noteworthy, IgG light chains, that are the second most represented nitroproteins in NPLD ultrafiltrates after albumin, are markedly increased in CKD serum being excreted and metabolized primarily by the kidney (Cohen and Horl 2009). These interfere with respiratory burst and apoptotic cell death mechanisms of neutrophils thus leading to defective cell-mediated immunity and impaired resolution of inflammation. The contribution of IgG nitration to IgG light chain accumulation and uremic toxicity remains unexplored.

As a key finding in this study, for the first time we have been able confirm by an unbiased approach that IgG chains are nitrated in vivo in uremic plasma. Using a chemoselective labeling procedure previously set up by some of us (Amoresano et al. 2007), nitropeptides corresponding to the lambda and heavy chains of uremic IgG were identified by bidimensional tandem mass spectrometry analysis and

nitration sites were specifically assigned. Further analyses are in progress to identify by means of this unbiased approach other nitrated proteins in plasma and a recent evolution of this procedure (Chiappetta et al. 2009) is expected to translate present qualitative data into quali-quantitative experiments.

In conclusion, with this study we provide the first extensive analysis of nitration targets in uremic plasma proteins carried out by immunorecognition and LC-MS/MS analysis. Since specificity of this qualitative investigation needs to be assessed for each protein target, in this study we developed a procedure of chemoselective labeling to obtain the unbiased identification of nitroproteins during bidimensional MS/MS of plasma proteins that are preliminarily screened and isolated by 2D PAGE and WB. This approach has been successfully applied for the first time to the identification of nitration sites in the human IgG molecule exposed to RNS in vivo, and we expect that it could be used in future large-scale proteomic studies to identify other plasma nitroproteins.

Moreover, the nitroproteome analysis of this study has revealed that uremic plasma shows an abnormal burden of nitration and the tentatively identified nitration targets suggest pathway and disease-specific events that deserve further investigation to be confirmed and to ascertain their clinical relevance with respect to immuno-inflammatory and cardiovascular symptoms of ESRD. Disease-specific differences in the relative abundance of nitrated forms cannot be excluded at present and will be further investigated using the chemoselective labeling and MS/MS analysis approach proposed in this study.

References

Abello N, Kerstjens HA, Postma DS, Bischoff R (2009) Protein tyrosine nitration: selectivity, physicochemical and biological consequences, denitration, and proteomics methods for the identification of tyrosine-nitrated proteins. J Proteome Res 8:3222–3238

Agalou S, Ahmed N, Babaei-Jadidi R, Dawnay A, Thornalley PJ (2005) Profound mishandling of protein glycation degradation products in uremia and dialysis. J Am Soc Nephrol 16:1471–1485

Amoresano A, Chiappetta G, Pucci P, D'Ischia M, Marino G (2007) Bidimensional tandem mass spectrometry for selective identification of nitration sites in proteins. Anal Chem 79:2109–2117

Buoncristiani U, Galli F, Benedetti S, Errico R, Beninati S, Ghibelli L, Floridi A, Canestrari F (1999) Quantitative and qualitative assessment and clinical meaning of molecules removed with BK membranes. Contrib Nephrol 125:133–158

Canavese C, Bergamo D, Ciccone G, Burdese M, Maddalena E, Barbieri S, Thea A, Fop F (2004) Low-dose continuous iron therapy leads to a positive iron balance and decreased serum transferrin levels in chronic haemodialysis patients. Nephrol Dial Transplant 19:1564–1570

Chiappetta G, Corbo C, Palmese A, Galli F, Piroddi M, Marino G, Amoresano A (2009) Quantitative identification of protein nitration sites. Proteomics 9:1524–1537

Cohen G, Horl WH (2009) Free immunoglobulin light chains as a risk factor in renal and extrarenal complications. Semin Dial 22:369–372

Eckardt KU, Berns JS, Rocco MV, Kasiske BL (2009) Definition and classification of CKD: the debate should be about patient prognosis—a position statement from KDOQI and KDIGO. Am J Kidney Dis 53:915–920

Galli F (2007) Protein damage and inflammation in uraemia and dialysis patients. Nephrol Dial Transplant 22(Suppl 5):v20–v36

Galli F, Benedetti S, Buoncristiani U, Piroddi M, Conte C, Canestrari F, Buoncristiani E, Floridi A (2003) The effect of PMMA-based protein-leaking dialyzers on plasma homocysteine levels. Kidney Int 64:748–755

Galli F, Benedetti S, Floridi A, Canestrari F, Piroddi M, Buoncristiani E, Buoncristiani U (2005) Glycoxidation and inflammatory markers in patients on treatment with PMMA-based protein-leaking dialyzers. Kidney Int 67:750–759

Gole MD, Souza JM, Choi I, Hertkorn C, Malcolm S, Foust RF 3rd, Finkel B, Lanken PN, Ischiropoulos H (2000) Plasma proteins modified by tyrosine nitration in acute respiratory distress syndrome. Am J Physiol Lung Cell Mol Physiol 278:L961–L967

Heffron SP, Parastatidis I, Cuchel M, Wolfe ML, Tadesse MG, Mohler ER 3rd, Ischiropoulos H, Rader DJ, Reilly MP (2009) Inflammation induces fibrinogen nitration in experimental human endotoxemia. Free Radic Biol Med 47:1140–1146

Himmelfarb J (2009) Uremic toxicity, oxidative stress, and hemodialysis as renal replacement therapy. Semin Dial 22:636–643

Ischiropoulos H (2009) Protein tyrosine nitration—an update. Arch Biochem Biophys 484:117–121

Jiao K, Mandapati S, Skipper PL, Tannenbaum SR, Wishnok JS (2001) Site-selective nitration of tyrosine in human serum albumin by peroxynitrite. Anal Biochem 293:43–52

Johnson RJ, Chenoweth DE (1985) Structure and function of human C5a anaphylatoxin. Selective modification of tyrosine 23 alters biological activity but not antigenicity. J Biol Chem 260:10339–10345

Koller H, Hochegger K, Zlabinger GJ, Lhotta K, Mayer G, Rosenkranz AR (2004) Apoptosis of human polymorphonuclear neutrophils accelerated by dialysis membranes via the activation of the complement system. Nephrol Dial Transplant 19:3104–3111

Kormoczi GF, Wolfel UM, Rosenkranz AR, Horl WH, Oberbauer R, Zlabinger GJ (2001) Serum proteins modified by neutrophil-derived oxidants as mediators of neutrophil stimulation. J Immunol 167:451–460

Matsuyama Y, Terawaki H, Terada T, Era S (2009) Albumin thiol oxidation and serum protein carbonyl formation are progressively enhanced with advancing stages of chronic kidney disease. Clin Exp Nephrol 13:308–315

McCall MN, Easterbrook-Smith SB (1989) Comparison of the role of tyrosine residues in human IgG and rabbit IgG in binding of complement subcomponent C1q. Biochem J 257:845–851

Mitrogianni Z, Barbouti A, Galaris D, Siamopoulos KC (2004) Tyrosine nitration in plasma proteins from patients undergoing hemodialysis. Am J Kidney Dis 44:286–292

Mitrogianni Z, Barbouti A, Galaris D, Siamopoulos KC (2009) Oxidative modification of albumin in predialysis, hemodialysis, and peritoneal dialysis patients. Nephron Clin Pract 113:c234–c240

National Kidney Foundation, Inc. (2002) KDOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification

Nowak P, Zbikowska HM, Ponczek M, Kolodziejczyk J, Wachowicz B (2007) Different vulnerability of fibrinogen subunits to

oxidative/nitrative modifications induced by peroxynitrite: functional consequences. Thromb Res 121:163–174
Ohmori H, Kanayama N (2005) Immunogenicity of an inflammation-associated product, tyrosine nitrated self-proteins. Autoimmun Rev 4:224–229
Parastatidis I, Thomson L, Fries DM, Moore RE, Tohyama J, Fu X, Hazen SL, Heijnen HF, Dennehy MK, Liebler DC et al (2007) Increased protein nitration burden in the atherosclerotic lesions and plasma of apolipoprotein A-I deficient mice. Circ Res 101:368–376
Parastatidis I, Thomson L, Burke A, Chernysh I, Nagaswami C, Visser J, Stamer S, Liebler DC, Koliakos G, Heijnen HF et al (2008) Fibrinogen beta-chain tyrosine nitration is a prothrombotic risk factor. J Biol Chem 283:33846–33853
Pawlak K, Pawlak D, Mysliwiec M (2006) Oxidative stress effects fibrinolytic system in dialysis uraemic patients. Thromb Res 117:517–522
Peluffo G, Radi R (2007) Biochemistry of protein tyrosine nitration in cardiovascular pathology. Cardiovasc Res 75:291–302
Pignatelli B, Li CQ, Boffetta P, Chen Q, Ahrens W, Nyberg F, Mukeria A, Bruske-Hohlfeld I, Fortes C, Constantinescu V et al (2001) Nitrated and oxidized plasma proteins in smokers and lung cancer patients. Cancer Res 61:778–784
Piroddi M, Depunzio I, Calabrese V, Mancuso C, Aisa CM, Binaglia L, Minelli A, Butterfield AD, Galli F (2007) Oxidatively-modified and glycated proteins as candidate pro-inflammatory toxins in uremia and dialysis patients. Amino Acids 32:573–592
Ptolemy AS, Lee R, Britz-McKibbin P (2007) Strategies for comprehensive analysis of amino acid biomarkers of oxidative stress. Amino Acids 33:3–18
Ryberg H, Caidahl K (2007) Chromatographic and mass spectrometric methods for quantitative determination of 3-nitrotyrosine in biological samples and their application to human samples. J Chromatogr B Analyt Technol Biomed Life Sci 851:160–171
Sezer MT, Akin H, Demir M, Erturk J, Aydin ZD, Savik E, Tunc N (2007) The effect of serum albumin level on iron-induced oxidative stress in chronic renal failure patients. J Nephrol 20:196–203
Shi XY, Hou FF, Niu HX, Wang GB, Xie D, Guo ZJ, Zhou ZM, Yang F, Tian JW, Zhang X (2008) Advanced oxidation protein products promote inflammation in diabetic kidney through activation of renal nicotinamide adenine dinucleotide phosphate oxidase. Endocrinology 149:1829–1839
Shishehbor MH, Aviles RJ, Brennan ML, Fu X, Goormastic M, Pearce GL, Gokce N, Keaney JF Jr, Penn MS, Sprecher DL et al (2003) Association of nitrotyrosine levels with cardiovascular disease and modulation by statin therapy. JAMA 289:1675–1680
Simmons EM, Langone A, Sezer MT, Vella JP, Recupero P, Morrow JD, Ikizler TA, Himmelfarb J (2005) Effect of renal transplantation on biomarkers of inflammation and oxidative stress in end-stage renal disease patients. Transplantation 79:914–919
Souza JM, Peluffo G, Radi R (2008) Protein tyrosine nitration—functional alteration or just a biomarker? Free Radic Biol Med 45:357–366
Sultana R, Boyd-Kimball D, Poon HF, Cai J, Pierce WM, Klein JB, Merchant M, Markesbery WR, Butterfield DA (2006) Redox proteomics identification of oxidized proteins in Alzheimer's disease hippocampus and cerebellum: an approach to understand pathological and biochemical alterations in AD. Neurobiol Aging 27:1564–1576
Thomson L, Christie J, Vadseth C, Lanken PN, Fu X, Hazen SL, Ischiropoulos H (2007) Identification of immunoglobulins that recognize 3-nitrotyrosine in patients with acute lung injury after major trauma. Am J Respir Cell Mol Biol 36:152–157
Vanholder R, Schepers E, Meert N, Lameire N (2006) What is uremia? Retention versus oxidation. Blood Purif 24:33–38
Ward RA (2005) Protein-leaking membranes for hemodialysis: a new class of membranes in search of an application? J Am Soc Nephrol 16:2421–2430
Wayenberg JL, Ransy V, Vermeylen D, Damis E, Bottari SP (2009) Nitrated plasma albumin as a marker of nitrative stress and neonatal encephalopathy in perinatal asphyxia. Free Radic Biol Med 47:975–982
Zheng L, Nukuna B, Brennan ML, Sun M, Goormastic M, Settle M, Schmitt D, Fu X, Thomson L, Fox PL et al (2004) Apolipoprotein A-I is a selective target for myeloperoxidase-catalyzed oxidation and functional impairment in subjects with cardiovascular disease. J Clin Invest 114:529–541

Research Article

Rapid Communications in Mass Spectrometry

Received: 28 July 2010 Revised: 15 October 2010 Accepted: 7 November 2010 Published online in Wiley Online Library: 00 Month 2010

Rapid Commun. Mass Spectrom. **2011**, *25*, 223–231
(wileyonlinelibrary.com) DOI: 10.1002/rcm.4863

Dansyl labeling and bidimensional mass spectrometry to investigate protein carbonylation

Angelo Palmese, Chiara De Rosa, Gennaro Marino and Angela Amoresano*

Department of Organic Chemistry and Biochemistry, School of Biotechnological Sciences, Federico II University of Naples, Naples, Italy

Carbonylation is a non-enzymatic irreversible post-translational modification. The adduction of carbonyl groups to proteins is due to the presence of excess of ROS in cells. Carbonylation of specific amino acid side chains is one of the most abundant consequences of oxidative stress; therefore, the determination of carbonyl groups content in proteins is regarded as a reliable way to estimate the cellular damage caused by oxidative stress. This paper reports a novel RIGhT (Reporter Ion Generating Tag) (A. Amoresano, G. Monti, C. Cirulli, G. Marino. *Rapid Commun. Mass Spectrom.* 2006, *20*, 1400) approach for selective labeling of carbonyl groups in proteins using dansylhydrazide, coupled with selective analysis by bidimensional mass spectrometry. We first applied this approach to ribonuclease A and lysozyme as model proteins. According to the so-called 'gel-free procedures', the analysis is carried out at the level of peptides following tryptic digest of the whole protein mixture. Modified RNaseA was analyzed in combined MS^2 and MS^3 scan mode, to specifically select the dansylated species taking advantage of the dansyl-specific fragmentation pathways. This combination allowed us to obtain a significant increase in signal/noise ratio and a significant increase in sensitivity of analysis, due to the reduction of duty cycle of the mass spectrometer. The unique signal obtained was correlated to peptide 1-10 of RNaseA carbonylated and labeled by dansylhydrazide. This strategy represents the first method leading to the direct identification of the carbonylation sites in proteins, thus indicating the feasibility of this strategy to investigate protein carbonylation in a proteomic approach. Copyright © 2010 John Wiley & Sons, Ltd.

Carbonylation is a non-enzymatic irreversible post-translational modification of proteins. The adduction of carbonyl groups to proteins is due to the presence of excess of ROS (Reactive Oxygen Species) in cells. These can react both with proteins and with other biological macromolecules like carbohydrates and lipids, generating more reactive by-products, namely RCS (Reactive Carbonyl Species). Proteins can react directly with ROS leading to the formation of derivatives possessing highly reactive carbonyl groups (ketones, aldehydes), or with reactive carbonyl compounds derived from oxidation of carbohydrates (glycoxidation products), lipids (malondialdehyde, 4-hydroxynonenal) and glycation and lipoxidation end products.[1,2]

Lysine, arginine, proline and threonine are particularly susceptible to direct oxidation on amino acid side chains leading to the formation of α-aminoadipic semialdehyde, glutamic semialdehyde, 2-pyrrolidone and 2-amino-3-ketobutyric acid, respectively.[3–5]

The increase in carbonylated proteins during cellular ageing or in response to oxidative stress is not a casual event; in fact some proteins are more susceptible to ROS than other. Therefore, the definition of the entire set of target proteins of oxidative stress is a hard challenge since these proteins are specie specific.[6–11]

* *Correspondence to*: A. Amoresano, Department of Organic Chemistry and Biochemistry, Federico II University of Naples, Monte S. Angelo, via Cynthia 4, 80126 Naples, Italy.
E-mail: angamor@unina.it

Carbonylation of specific amino acid side chains[3–5] is one of the most abundant consequences of oxidative stress; so determination of the content of carbonyl groups in proteins is regarded as a reliable way to estimate the cellular damage caused by oxidative stress. For this reason several methods have been developed to determine the content of carbonyl groups in proteins.[12–17] A simple and accurate method to quantify protein oxidation uses 2,4-dinitrophenylhydrazine, which specifically reacts with carbonyls to form 2,4-dinitrophenylhydrazone;[18,19] the amount of hydrazone is quantified by using a spectrophotometer technique. However, this method is not suitable for proteins containing chromophores adsorbing at 360–380 nm (i.e. heme-containing proteins) and requires large amounts of proteins, not always available. The DNPH assay has also been used in combination with Western blot analysis using anti-dinitrophenyl antibodies, enhancing the sensitivity of the method.[20–22]

As an alternative, fluorescent methods have been developed based on the use of fluorescein hydrazyde[23] that reacts specifically with carbonyl groups and can be detected fluorometrically at 489 nm.

A radioactive method has been developed that use sodium borotritide for reduction of carbonyl groups and their Schiff bases; this methodology has some drawbacks such as the use of radioactive reagents with risks for safety and non-specific incorporation of tritium into proteins.[24]

A proposed method for detection of carbonyl groups uses biotinhydrazide for carbonyl labeling; labeled proteins can be analyzed exploiting the specific interaction biotin-avidin; this strategy has been used in combination with an affinity

chromatography step and mass spectrometric analysis to identify specific sites of carbonylation in proteins.[25–31]

We have recently reported an innovative approach, named RIGhT (Reporter Ion Generating Tag), which involves dansyl labeling of nitration and phosphorylation sites and relies on the enormous potential of MS^n analysis.[32,33] Taking advantage of the experience already made, this work reports a novel RIGhT methodology for selective labeling of carbonyl groups in proteins using dansylhydrazide coupled with selective analysis in precursor ion scanning mode.

EXPERIMENTAL

Chemicals

Dithiothreitol (DTT), trypsin, α-cyano-4-hydroxycinnamic acid, ethylenediaminetetraacetic acid (EDTA), urea, sodium hypochlorite solution and bovine ribonuclease A (RNaseA) were from Sigma (St. Louis, MO, USA). 5-*N*,*N*-(Dimethylamino)naphthalene-1-sulfonohydrazide (DNSH), ammonium hydrogen carbonate (AMBIC) and iodoacetamide (IAM) were purchased from Fluka. Acetonitrile (ACN) was from Romil. Trifluoroacetic acid (TFA) (HPLC grade) was purchased from Carlo Erba. All solvents were of the highest purity available from Baker. All other reagents and proteins were of the highest purity available from Sigma.

In vitro carbonylation of RNaseA

RNaseA was dissolved in a 5 mM NaClO solution to a final concentration of 0.5 mg/mL. The reaction was carried out at 37°C for 15 min. Carbonylated RNaseA was immediately desalted by size exclusion chromatography (SEC) on a Sephadex G-25M column (GE Healthcare). Protein elution was monitored at 220 and 280 nm. The fraction containing proteins was collected, lyophilized and stored at −20°C.

Disulfide bridge reduction, Cys alkylation and tryptic digestion

Carbonylated RNaseA was dissolved in denaturant buffer (urea 6 M, Tris 300 mM, EDTA 10 mM pH 8.0) and disulfide bridges were reduced with DTT (10-fold molar excess on the Cys residues) at 37°C for 2 h and then alkylated by adding IAM (5-fold molar excess on thiol residues) at room temperature for 30 min in the dark. Protein samples were desalted by SEC on a Sephadex G-25M column (GE Healthcare). The fraction containing proteins was lyophilized and then dissolved in 10 mM AMBIC buffer pH 8.0. Trypsin digestion was performed using an enzyme/substrate ratio of 1:50 (w/w) at 37°C for 16 h.

Bacterial strains, growth conditions and protein extract preparation

Escherichia coli K12 strain was grown in aerobic conditions at 37°C in LB medium. After 16 h, bacteria were harvested by centrifugation and resuspended in Buffer Z (25 mM HEPES pH 7.6, 50 mM KCl, 12.5 mM $MgCl_2$, 1 mM DTT, 20% glycerol, 0.1% Triton) containing 1 mM PMSF. Cells were disrupted by sonication. The suspension was centrifuged at 9000 *g* for 30 min at 47°C. After centrifugation the protein concentration of the extract was determined by Bradford assay.

Labeling of carbonyl groups with DNSH

The sample was lyophilized and then dissolved in 10 μL of H_2O. To this solution were added 100 μL of a 6.5% TFA solution in ACN and an equal volume of a 5 mM DNSH solution in ACN. Solutions were allowed to react at 37°C for 8 h.

Mass spectrometry

Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) experiments were performed on a Voyager-DE STR MALDITOF mass spectrometer (Applied Biosystems, Framingham, MA, USA) equipped with a nitrogen laser (337nm). A 1 μL volume of the total mixture was mixed (1:1, v/v) with a 10 mg/ml solution of α-cyano-4-hydroxycinnamic acid in ACN/50 mM citrate buffer (70:30 v/v).

Nanospray liquid chromatography/tandem mass spectrometry (nanoLC-MS/MS) experiments were performed on a 4000 Q-Trap mass spectrometer (Applied Biosystems) coupled to a 1100 nanoHPLC system (Agilent Technologies). Peptide mixtures were loaded onto an Agilent reversed-phase pre-column cartridge (Zorbax 300 SB-C18, 5 × 75 μm, 3.5 μm) at 20 μL/min with solvent A (0.1% formic acid, 2% ACN in water, loading time 5 min). Peptides were then separated on an Agilent reversed-phase column (Zorbax 300 SB-C18, 150 mm × 75 μm, 3.5 μm), at a flow rate of 300 nL/min using 0.1% formic acid, 2% ACN in water as solvent A and 0.1% formic acid, 2% water in ACN as solvent B. The elution was accomplished by a 7–50% linear gradient of solvent B in 50 min. A micro-ionspray source was used at 2 kV with liquid coupling, with a declustering potential of 50 V, using a uncoated silica tip (o.d. 150 μm, i.d. 20 μm, tip diameter 10 μm; New Objectives, Ringoes, NJ, USA). Spectra acquisition was based on a survey precursor ion scan for *m/z* 170. The Q_1 was scanned from 400 to 1400 in 2 s with resolution 'low', and the precursor ions were fragmented in q_2 using a linear gradient of collisional energy from 30 to 80 V. Finally, Q_3 was set to transmit only ions at *m/z* 170 with resolution 'unit'. This scan mode was followed by an enhanced resolution experiment for the ions of interest and then by MS^2 and MS^3 experiments of the two most abundant species. MS^2 spectra were acquired using the best collision energy calculated on the basis of *m/z* values and charge state (*rolling collision energy*). MS^3 experiments were performed on the fragment ion at *m/z* 234, with a fixed linear ion trap (LIT) fill time of 350 ms and an activation time of 100 ms, scanning the mass range from 160 to 180 *m/z*.

RESULTS AND DISCUSSION

The analytical strategy described in this paper fulfils the proteomics needs of sensitivity and specificity in the analysis of substoichiometric post-translational modifications, such as protein carbonylation, associated with protein oxidation.[33] Existing methodologies, essentially based on immunochemical techniques,[20–22] which although sensitive, lack the capability to precisely localize modification sites, information that is often required to completely understand the biological significance of modification. Here we suggest a derivatization procedure based on dansylhydrazide having a

chemical moiety that gives rise to characteristic product ions, thus allowing the exploitation of tandem multi-stage mass spectrometric (MS^n) analysis. As exploited before, 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 highly reproducible.[31,32] We tested this approach using ribonuclease A and lysozyme as model proteins. According to 'gel-free procedures', the analysis is carried out at the level of peptides following tryptic digest of the whole protein mixture. RNaseA (0.5 mg) was carbonylated with NaClO under controlled conditions,[14] and then rapidly desalted by SEC as described in the Experimental section. NaClO-induced modifications were then analyzed by MALDI mapping analysis. Aliquots of the sample containing RNaseA and carbonylated RNaseA were reduced with DTT and alkylated with IAM. The protein samples were digested with trypsin, and the resulting peptides mixtures analyzed by MALDI-MS. Mass spectral analyses led to the verification of some 95% sequence coverage and allowed us to locate the carbonylated residue in the peptide 1-10. In fact, MALDI-MS spectra show the occurrence of a signal at *m/z* 1149.6 (Figs. 1(A) and 1(B)) exhibiting a molecular mass not corresponding to a theoretical one within the RNaseA sequence. This signal was assigned to the peptide 1-10 showing a $\Delta M = -1$ Da. This value corresponds to the mass difference existing between lysine and α-aminoadipic semialdehyde after a protein carbonylation event. In fact, protein carbonyls generated by oxidation of lysine, arginine, or proline residues, and frequently used as markers for oxidative protein modification, determine a difference in the molecular mass of amino acid side chains.[4] Oxidation of the proline or arginine side chain forms glutamic semialdehyde, while lysine oxidation results in aminoadipic semialdehyde, introducing carbonyl groups in the protein structure. The hydroxyl group of the threonine side chain can

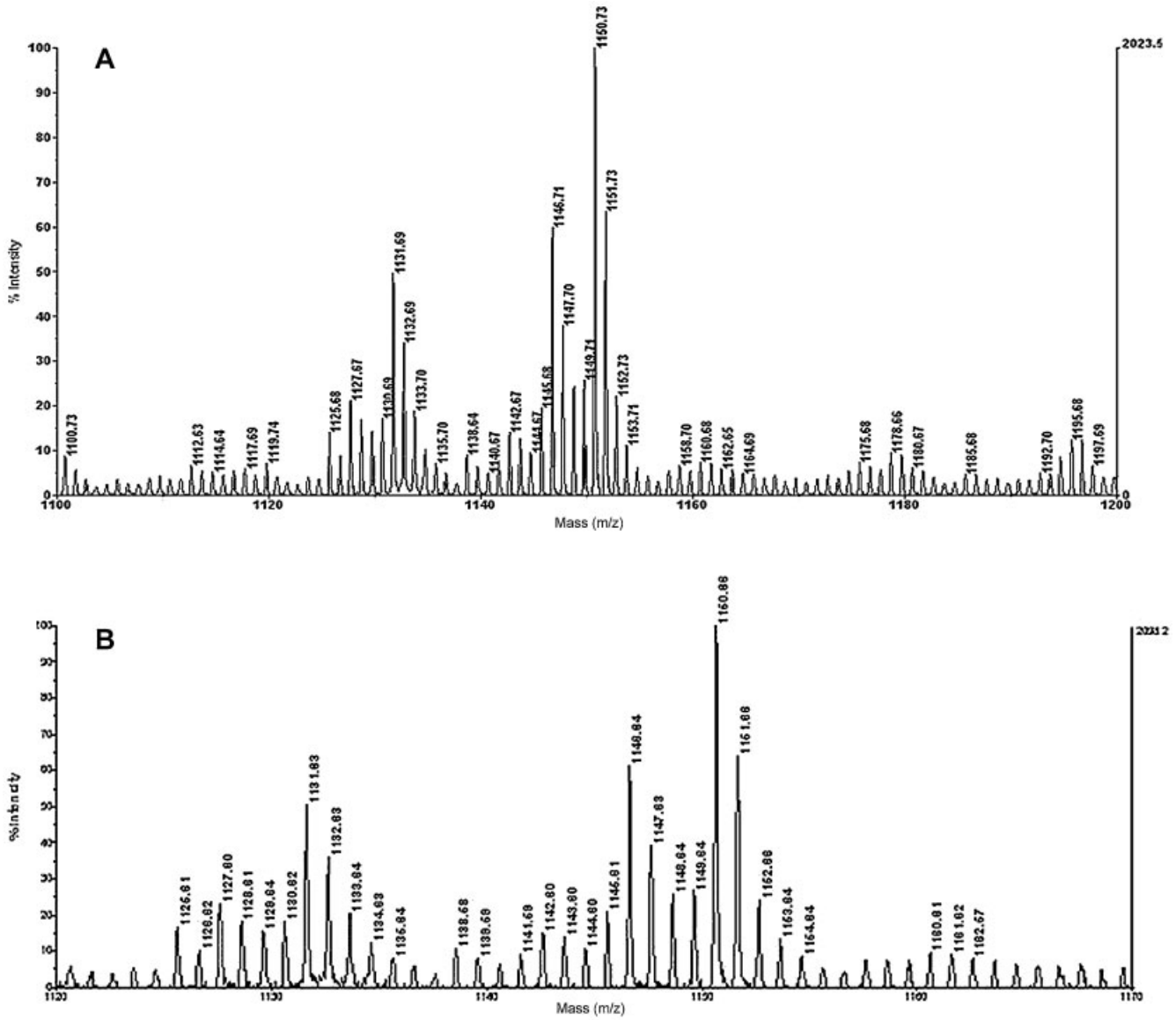


Figure 1. (A) Partial MALDI-MS spectrum shows the occurrence of a signal at *m/z* 1149.6 attributed to the peptide 1-10 of RNaseA with $\Delta M = -1$ Da, corresponding to the mass difference existing between lysine and α-aminoadipic semialdehyde after a protein carbonylation event. (B) Zoomed region of the MALDI-MS spectrum of carbonylated RNaseA.

also be oxidized to form the carbonyl group. The presence in the MALDI spectrum of the signal at *m/z* 1150.6 attributed to the theoretical peptide 1-10 indicated that the carbonylation reaction was not quantitative. No signal occurring at *m/z* 1148.6, and corresponding to the peptide 1-10 doubly modified by carbonylation at the second residue of lysine, was detected. However, the mass spectrum showed the occurrence of a signal at *m/z* 1131.6 (Figs. 1(A) and 1(B)) showing a $\Delta M = -18$ Da with respect to the carbonylated peptide. This species was probably due to the formation of a cyclic intermediate originating from the reaction of the N-terminal amino group with the aldehyde of the oxidized lysine residue in position 1. A similar procedure was applied to lysozyme carbonylation. The tryptic mixture of carbonylated lysozyme was then submitted to MALDI-MS analyses. Interpretation of MALDI-MS spectra showed the occurrence of a signal whose molecular mass could not be explained on the basis of the theoretical ones. It should be noted that, like RNaseA, lysozyme showed the presence of a lysine residue at the N-terminal position. As detected before, the signal at *m/z* 605.3 corresponding to carbonylated peptide 1-5 was not detected. However, a signal at *m/z* 587.3 (18 Da lower than carbonylated species) was detected. This peak was absent in the spectrum of wild-type lysozyme (Fig. 2(A)), and was tentatively assigned to the cyclic intermediate. To support our hypothesis, we decided to modify the strategy, introducing an acetylation step to block amino groups (namely N-terminus and Lys ε-amino groups).[34] Following the

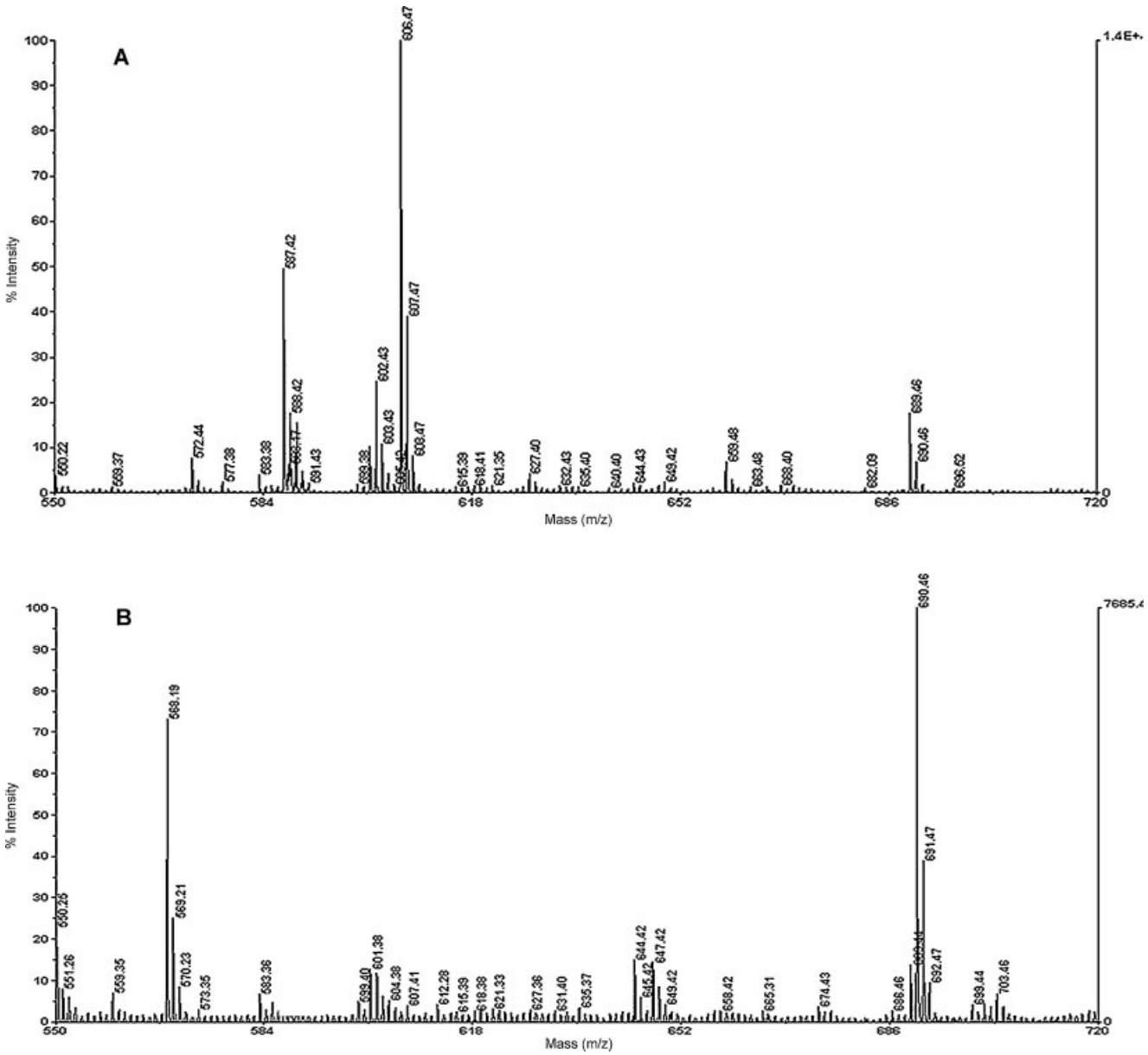


Figure 2. Partial MALDI-MS spectra of carbonylated lysozyme (A) and acetylated carbonylated lysozyme (B). In (A) the signal at *m/z* 606.4 is attributed to the peptide 1-5 of lysozyme, the signal at *m/z* 587.4 is tentatively assigned to a cyclic intermediate formed after carbonylation of the protein. In (B) two related signals occurred at *m/z* 647.3 and 690.3. These signals were attributed to the peptide 1-5 carbonylated and acetylated (3M = + 42 Da) and the same peptide doubly acetylated, respectively. The absence of the signal at *m/z* 587.3 could be appreciated, thus confirming the formation of the cyclic intermediate at the N-terminal residue.

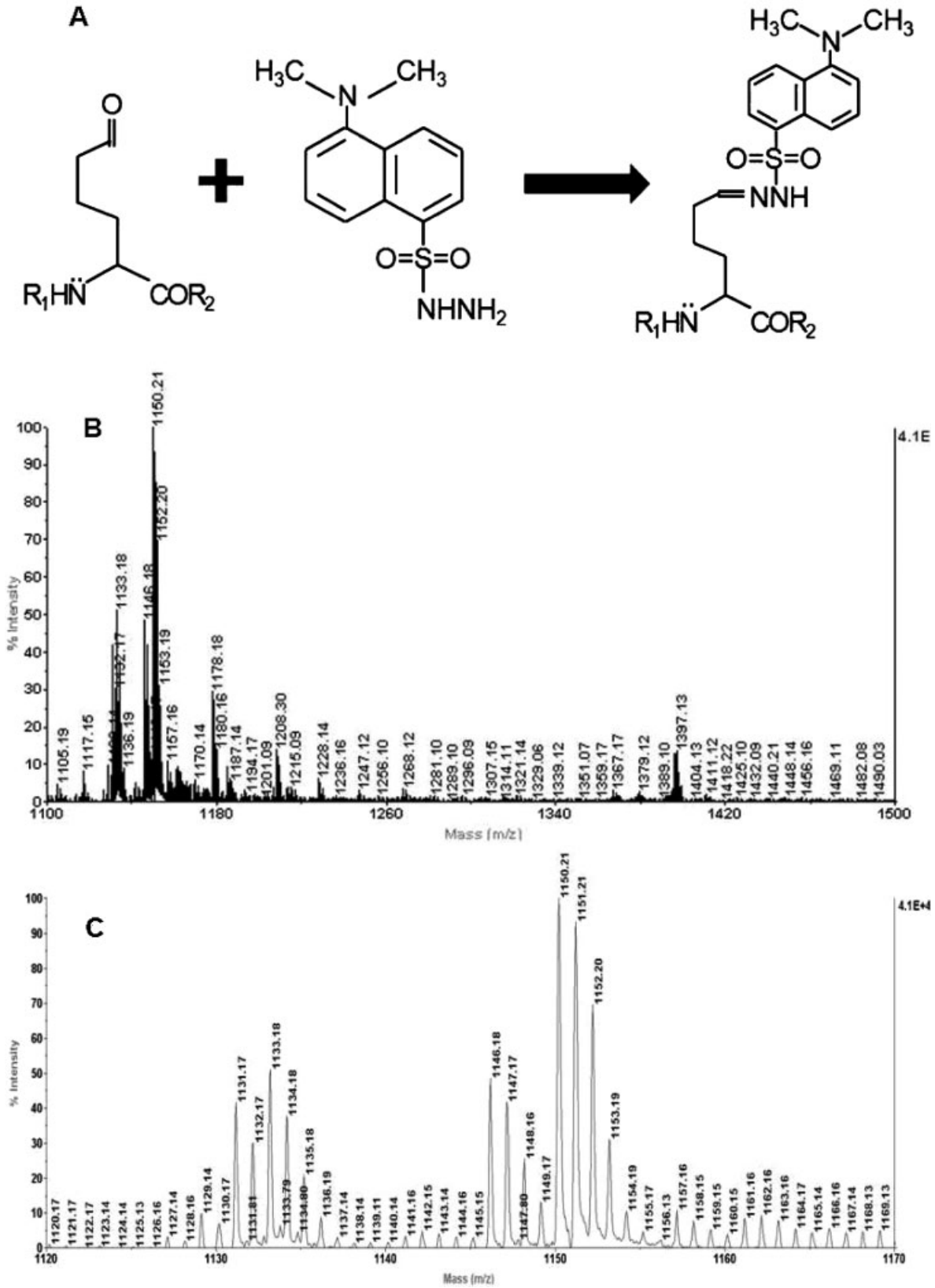


Figure 3. (A) Formation of the α-aminoadipic semialdehyde hydrazone following dansylhydrazide reaction. (B) MALDI-MS spectra of the carbonylated RNaseA peptide mixture after dansylhydrazide reaction. The signal at *m/z* 1396.7 was detected showing a mass increment of +247 Da attributed to the peptide 1-10 modified by hydrazone formation (ΔM = +247 Da). (C) Zoomed region of the MALDI-MS spectrum.

reaction, the acetylated N-terminal amino group cannot react with the aldehyde moiety occurring at the oxidized residue. The acetylation reaction was carried out by using acetic acid *N*-hydroxysuccinimide ester (1000-fold molar excess on proteins). The mixtures were then analyzed by MALDI-TOF MS and the spectra revealed the disappearance of signals attributed to the cyclic intermediate formation. As an example the MALDI-MS spectrum recorded from the acetylated and oxidized tryptic peptide mixture from lysozyme showed the presence of two related signals at *m/z* 647.3 and 690.3. These signals were attributed to the peptide 1-5 carbonylated and acetylated ($\Delta M = +42$ Da) and the same peptide doubly acetylated, respectively. The absence of the signal at *m/z* 587.3 (Fig. 2(B)) could be seen, thus confirming the formation of the cyclic intermediate when α-aminoadipic semialdehyde is at the N-terminal residue. The formation of the cyclic intermediate blocks the reactivity of carbonylated lysine, which is not suitable for reaction with the hydrazide moiety of the DNS derivative.

Labeling with DNSH

Because of the non-reactivity of the cyclic intermediate we decided to work on carbonylated RNaseA only. Moreover, the acetylation of carbonylated lysozyme means the introduction of a further step that complicates the analytical strategy and is not useful for our purpose. Labeling of carbonylated residues of RNaseA was carried out using dansylhydrazide. The yield of the addition reaction was monitored by MALDI-MS analysis. Mass spectral analysis showed the decrease of intensity of the signal at *m/z* 1149.6 (Fig. 3(C)), indicating the labeling of this species. A new signal at *m/z* 1396.7 was detected showing a mass increment of +247 Da (Fig. 3(B)). This signal was attributed to the peptide 1-10 modified by α-aminoadipic semialdehyde hydrazone formation following dansylhydrazide reaction (Fig. 3(A)) on the RNaseA. Moreover, the absence of further species with a $\Delta M = +247$ Da probably indicates the specificity of the reagent for carbonyls, i.e. carbonylated lysine. The MALDI-MS spectrum still shows the occurrence of the signal at *m/z* 1131.6 assigned to the cyclic intermediate originating from the reaction of the N-terminal amino

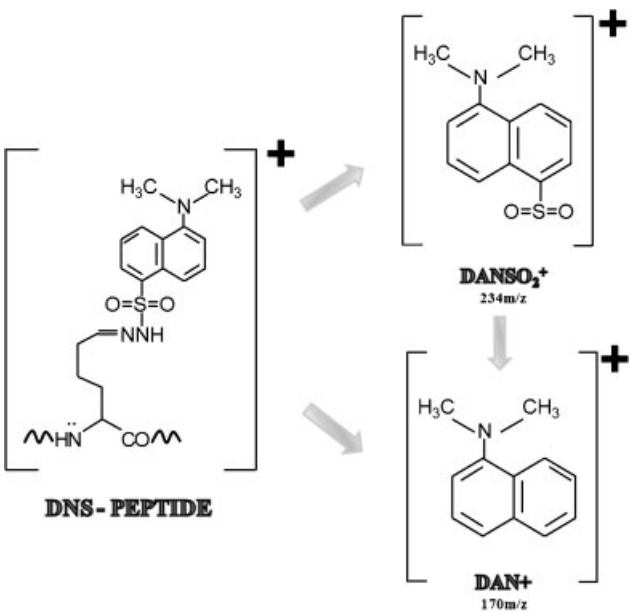


Figure 4. Fragmentation pathways of dansyl derivatives after collision-induced dissociation. The specific fragmentation can be used for selective precursor ion scan analysis for the fragment at *m/z* 170 and the transition *m/z* 234→170 can be used as dansyl-specific MS3 transition.

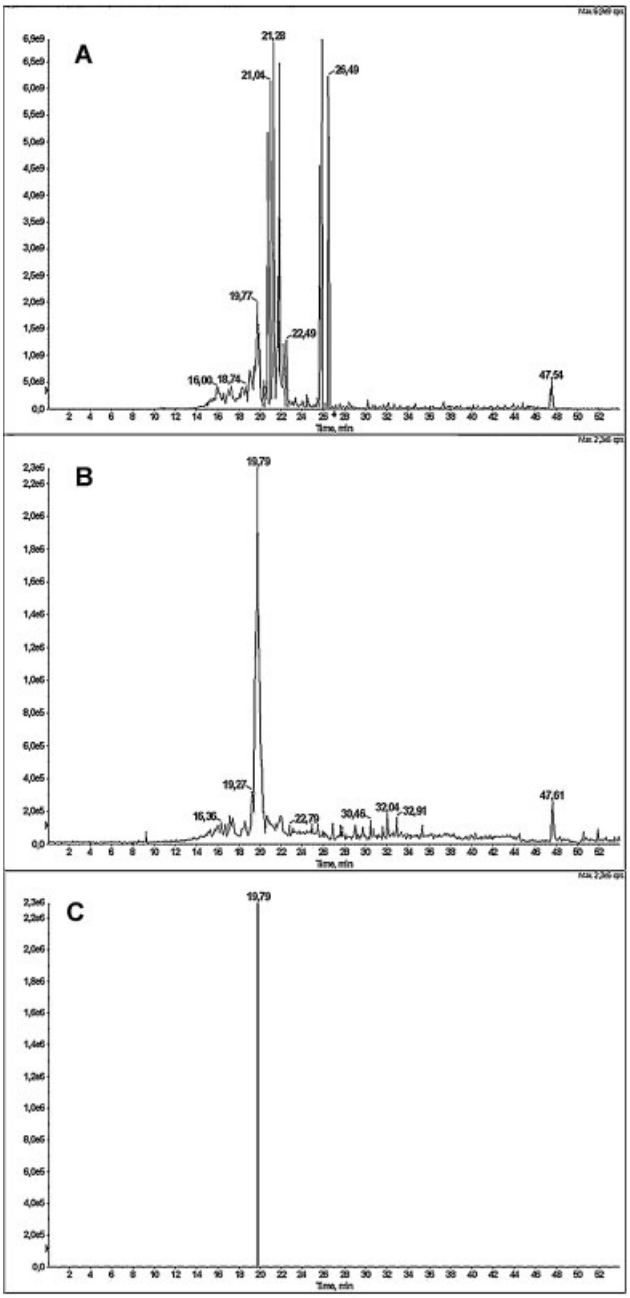


Figure 5. LC/MS/MS chromatograms of dansylated RNase A trypsin mixture: (A) TIC of the full scan MS/MS analysis, in which a large number of signals most of which are not related to dansyl-modified peptides are present. The combination of PIS and MS^3 (B, C) analysis allowed us to obtain a significant increase in signal/noise ratio and a significant increase in sensitivity of the mass spectrometric analysis. The unique signal at 19.8 min in the PIS TIC was correlated to peptide 1-10 of RNase A carbonylated and labeled by dansylhydrazide.

group with the aldehyde of the oxidized lysine residue in position 1. The formation of this intermediate determines the disappearance of carbonyl groups of lysine in position 1, thus this residue is not yet reactive towards the hydrazide moiety.

LC/MS/MS analysis of dansylated peptides

Discrimination between modified and non-modified peptides is achieved by taking advantage of the instrumental features of a hybrid (triple quadrupole/linear ion trap) mass spectrometer. Peptide analysis is carried out by LC/MS/MS, and the peptide ions of interest are discriminated by two selectivity criteria based on two subsequent tandem MS experiments, a precursor ion scan followed by an MS^3 scan mode in a bidimensional tandem mass spectrometry selection.[33] Thus to locate the modification site on the peptide, we performed an LC/MS/MS analysis of the entire peptide mixture using a linear ion trap 4000 Q-Trap instrument, operating with *Precursor Ion Scan* as survey scan mode. The precursor ion scan mode allowed us to analyze the only precursor ions producing the fragment at *m/z* 170, generated by typical fragmentation of dansyl and its derivatives.[31] The selected precursor ions were then analyzed in combined MS^2 and MS^3 scan mode, to specifically select the dansylated species by combining the Precursor Ion Scan analysis for the fragment at *m/z* 170 and taking advantage of the dansyl-specific MS^3 transition *m/z* 234→170 (Fig. 4). Figure 5 shows LC/MS/MS chromatograms from the dansylated RNaseA trypsin mixture. As clearly indicated in the figure, by comparison of the TIC profiles from PIS and MS^3 (Figs. 5(B) and 5(C)) with the TIC chromatogram of LC/MS/MS full scan analysis (Fig. 5(A)), the precursor ion scan mode still showed the occurrence of a large number of signals, most of which not related to dansyl-modified peptides. In fact, the corresponding MS^2 spectra do not show the occurrence of a fragment ion at *m/z* 234. Further selection based on the MS^3 scan removed a large number of false positives leading to a simple ion chromatogram essentially dominated by a unique signal, indicating the presence of a single dansyl-modified peptide. The combination of PIS and MS^3 analysis allowed us to obtain a significant increase in signal/noise ratio and a significant increase in the sensitivity of the mass spectrometric analysis, due to the reduction of the duty cycle of the mass spectrometer. The unique signal at 19.8 min in the PIS TIC was correlated by its MS and MS/MS spectra to peptide 1-10 of RNaseA carbonylated and labeled by dansylhydrazide (Fig. 5). Moreover, the modified lysine residue was easily detected by interpretation of the MS spectrum and thus confirming the previous hypothesis of carbonylation at the level of Lys 7 within RNaseA. Figure 6 shows the MS^2 spectrum of the triply charged ion at *m/z* 466.2 of peptide 1-10 carrying a dansylhydrazone moiety. The MS^2 spectrum related to the peptide 1-10 showed the presence of fragment ions belonging to both the y and the b series. However, some features make interpretation of MS/MS spectra problematical. In fact the modified peptide is not very stable during collision-induced dissociation, so that the peptide bond between alanine in position 6 and modified α-aminoadipic semialdehyde in position 7 was not completely fragmented under the conditions used in the experiment. Thus the occurrence of a signal corresponding to the y_4 ion can be detected even at very low intensity. This fragment ion in fact corresponds to the peptide sequence KFER still containing Lys modified by a dansyl moiety. Moreover, the DNS neutral loss (*m/z* 234 Da) from the modified α-aminoadipic semialdehyde residue can be appreciated. In fact, y ions exhibiting a ΔM of 233 Da, corresponding to y fragments that have lost the DNS moiety but still retaining the hydrazone moiety, can be observed in the spectrum (Fig. 6).

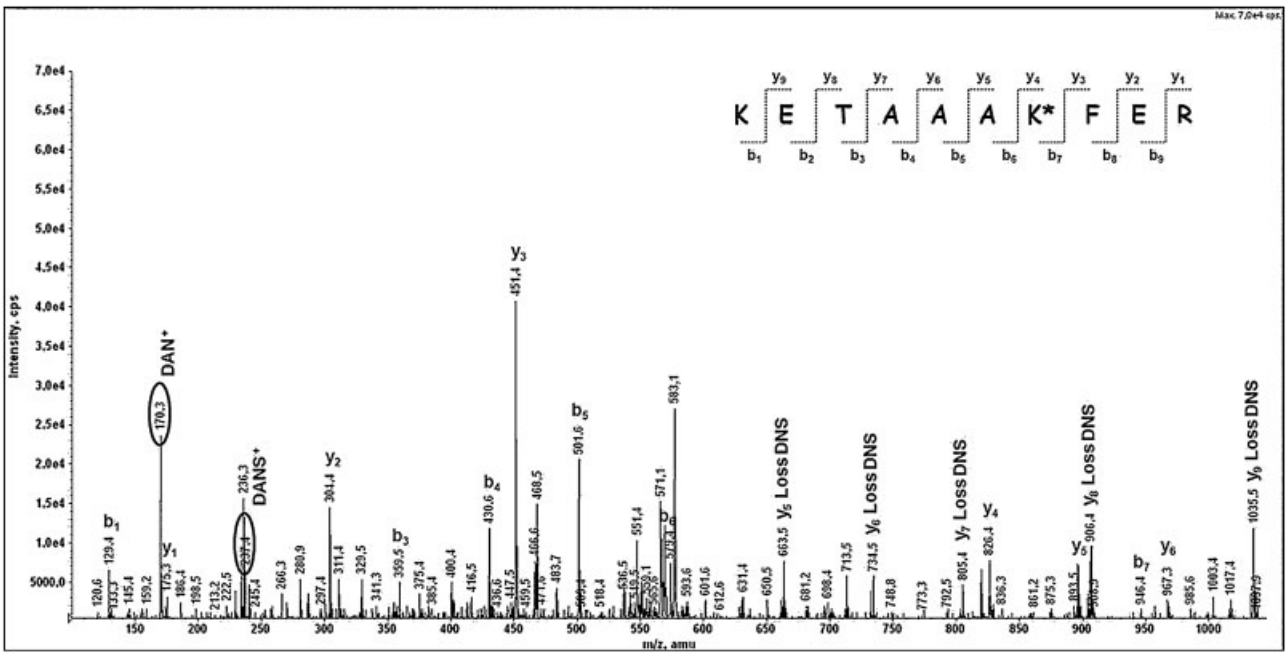


Figure 6. MS^2 spectrum of triply charged ion at *m/z* 466.2 of the peptide 1-10 carrying a dansylhydrazone moiety. The attributions are properly annotated.

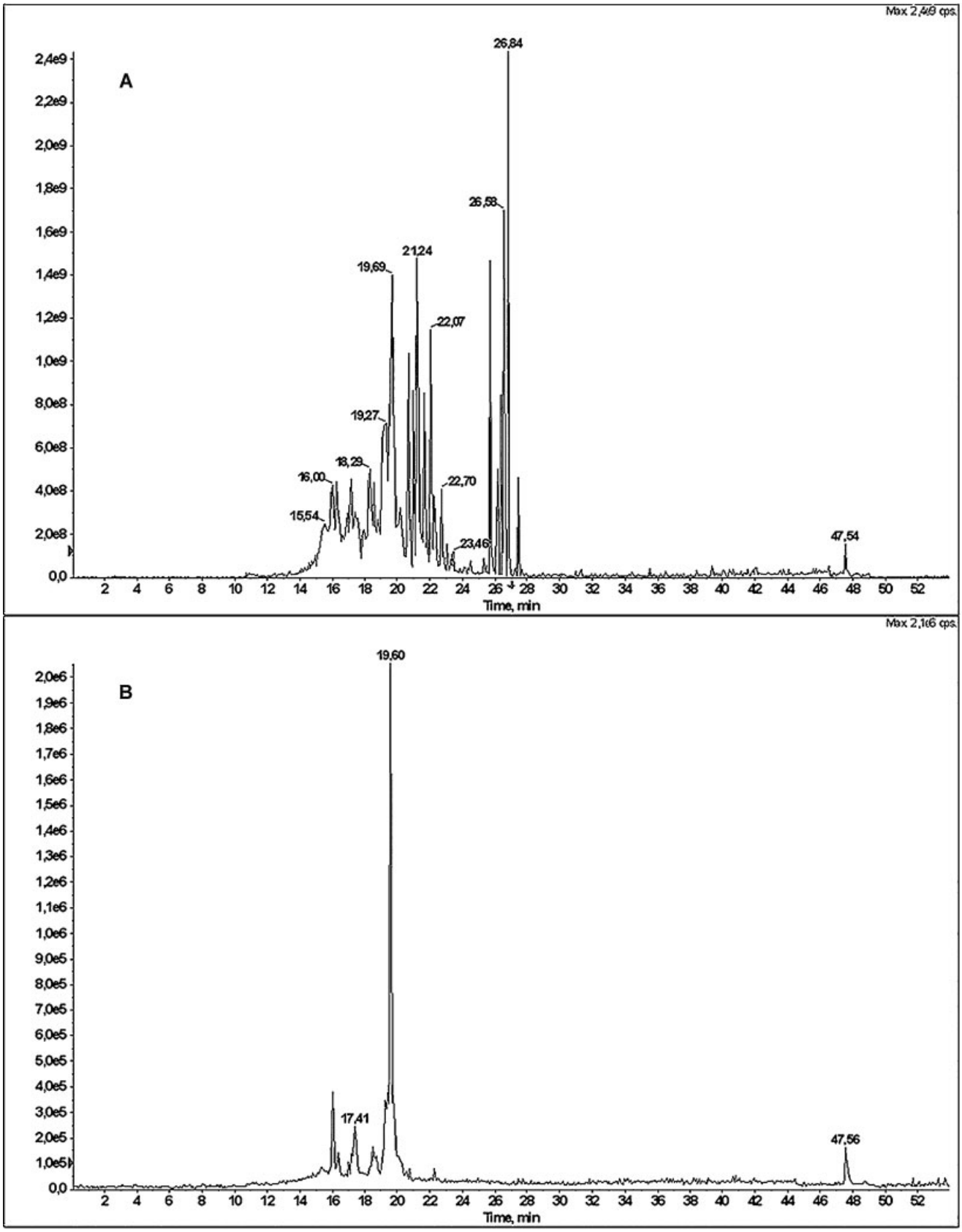


Figure 7. Chromatographic profiles of TIC (A) and PIS (B) experiments obtained by bidimensional LC/MS/MS analysis of 100 μg of the carbonylated RNase A spiked to 10 mg of entire cellular extract of *Escherichia coli*.

Analysis of a complex protein mixture

The proposed strategy was then confronted with a more complex sample, in order to demonstrate the ability of this analytical method to specifically locate carbonylated residues in a proteomics analysis. Carbonylated RNaseA (100 μg) was spiked to 10 mg of entire cellular extract of *Escherichia coli*. This sample was then analyzed using the procedure described above. Proteins were reduced by DTT and thiols were alkylated with IAM; then the protein mixture was digested with trypsin and the resulting peptide mixture was submitted to labeling with dansylhydrazide as described in the Experimental section. An aliquot of the resulting dansylated peptide mixture was submitted to bidimensional LC/MS/MS analysis in PIS and MS^3 modes as described above. Results of LC/MS/MS analysis are summarized in Fig. 7, showing the chromatographic profiles of TIC (Fig. 7(A)) and PIS (Fig. 7(B)) experiments. In the TIC profile a large number of signals are present, that could be due both to false positives or real dansylated peptides. However, the selectivity of the bidimensional analysis allowed us to eliminate all false positives since a unique signal referable to carbonylated and dansylated peptide 1-10 of RNaseA was detected. The manual interpretation of the MS^2 spectrum related to this signal allowed us to reconstruct the above mentioned peptide and to locate the modification site at the level of the lysine in position 7 within the peptide 1-10. To the best of our knowledge this strategy represents the first method leading to the direct identification of carbonylation sites in proteins, thus indicating the feasibility of this strategy to investigate protein carbonylation in a proteomic approach (manuscript in preparation).

CONCLUSIONS

In this study the use of dansylhydrazide in combination with advanced mass spectrometric experiments allowed us to specifically identify carbonylation sites of RNaseA also in a complex mixture as the entire protein extract of *E. coli*. The generation of carbonyls group in RNaseA was obtained by the oxidation with NaClO, with a slight modification of the protocol proposed by Headlam and Davies,[14] and allowed us to specifically convert just one lysine residue (namely Lys 7) of RNaseA into α-aminoadipic semialdehyde. Dansylhydrazide was then used to specifically label the α-aminoadipic semialdehyde residue and the peculiar fragmentation pathway of dansyl derivatives was then exploited for mass spectrometric analysis with a hybrid instrument, such as a triple quadrupole/linear ion trap mass spectrometer, in *Precursor Ion Scan* analysis combined with MS^3 analysis, in a bidimensional LC/MS/MS analysis. The combination of PIS and MS^3 analysis allowed us to obtain a significant increase in signal/noise ratio and a significant increase in sensitivity of mass spectrometric analysis, due to the reduction of the duty cycle of the mass spectrometer. In this way we were able to specifically identify the only modified peptide of RNaseA and also to locate the modification sites, namely the oxidized residue. This work represents the first method leading to the direct identification of the carbonylation sites in proteins, thus indicating the feasibility of this strategy to investigate protein carbonylation in a proteomic approach.

Acknowledgements

This work was supported by Italian MIUR grants, PRIN 2007, and FIRB Italian Human ProteomeNet Project 2007.

REFERENCES

[1] P. A. Grimsrud, H. Xie, T. J. Griffin, D. A. Bernlohr. *J. Biol. Chem.* **2008**, *283*, 21837.
[2] K. J. Davies, M. E. Delsignore, S. W. Lin. *J. Biol. Chem.* **1987**, *262*, 9902.
[3] E. R. Stadtman, R. L. Levine. *Amino Acids* **2003**, *25*, 207.
[4] J. R. Requena, C. C. Chao, R. L. Levine, E. R. Stadtman. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 69.
[5] J. R. Requena, R. L. Levine, E. R. Stadtman. *Amino Acids* **2003**, *25*, 221.
[6] B. S. Yoo, F. E. Regnier. *Electrophoresis* **2004**, *25*, 1334.
[7] B. Magi, A. Ettorre, S. Liberatori, L. Bini, M. Andreassi, S. Frosali, P. Neri, V. Pallini, A. Di Stefano. *Cell Death Differ.* **2004**, *11*, 842.
[8] E. Johansson, O. Olsson, T. Nyström. *J. Biol. Chem.* **2004**, *279*, 22204.
[9] K. England, C. O'Driscoll, T. G. Cotter. *Cell Death Differ.* **2004**, *11*, 252.
[10] K. England, T. Cotter. *Biochem. Biophys. Res. Commun.* **2004**, *320*, 123.
[11] B. A. Soreghan, F. Yang, S. N. Thomas, J. Hsu, A. J. Yang. *Pharm Res.* **2003**, *20*, 1713.
[12] R. L. Levine, N. Wehr, J. A. Williams, E. R. Stadtman, E. Shacter. *Methods Mol. Biol.* **2000**, *99*, 15.
[13] P. Ghezzi, V. Bonetto. *Proteomics* **2003**, *3*, 1145.
[14] H. A. Headlam, M. J. Davies. *Free Radical Biol. Med.* **2004**, *36*, 1175.
[15] A. G. Madian, F. E. Regnier. *J Proteome Res.* **2010**, *9*, 1330.
[16] H. Mirzaei, F. Regnier. *Anal. Chem.* **2005**, *77*, 2386.
[17] H. Mirzaei, F. Regnier. *Anal. Chem.* **2006**, *78*, 770.
[18] R. L. Levine, D. Garland, C. N. Oliver, A. Amici, I. Climent, A. G. Lenz, B. W. Ahn, S. Shaltiel, E. R. Stadtman. *Methods Enzymol.* **1990**, *186*, 464.
[19] R. L. Levine, J. A. Williams, E. R. Stadtman, E. Shacter. *Methods Enzymol.* **1994**, *233*, 346.
[20] A. Nakamura, S. Goto. *J. Biochem.* **1996**, *119*, 768.
[21] R. Sharma, A. Nakamura, R. Takahashi, H. Nakamoto, S. Goto. *Free Radical Biol. Med.* **2006**, *40*, 1179.
[22] A. C. Blackburn, W. F. Doe, G. D. Buffinton. *Free Radical Biol Med.* **1999**, *27*, 262.
[23] B. Ahn, S. G. Rhee, E. R. Stadtman. *Anal. Biochem.* **1987**, *161*, 245.
[24] A. G. Lenz, U. Costabel, S. Shaltiel, R. L. Levine. *Anal. Biochem.* **1989**, *177*, 419.
[25] A. Temple, T. Y. Yen, S. Gronert. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1172.
[26] H. Mirzaei, F. Regnier. *J. Chromatogr. A* **2007**, *1141*, 22.
[27] D. L. Meany, H. Xie, L. V. Thompson, E. A. Arriaga, T. J. Griffin. *Proteomics* **2007**, *7*, 1150.
[28] H. Mirzaei, B. Baena, C. Barbas, F. Regnier. *Proteomics* **2008**, *8*, 1516.
[29] K. Hensley. *Methods Mol. Biol.* **2009**, *536*, 457.
[30] T. J. Caperna, A. E. Shannon, A. le Blomberg, W. M. Garrett, T. G. Ramsay. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **2010**, *156*, 189.
[31] A. Amoresano, G. Monti, C. Cirulli, G. Marino. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 1400.
[32] A. Amoresano, G. Chiappetta, P. Pucci, M. D'Ischia, G. Marino. *Anal. Chem.* **2007**, *79*, 2109.
[33] A. Nakamura, K. Kawakami, F. Kametani, H. Nakamoto, S. Goto. *Ann. NY Acad. Sci.* **2010**, *1197*, 33.
[34] F. L. Brancia, S. G. Oliver, S. J. Gaskell. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 2070.

Appl Microbiol Biotechnol
DOI 10.1007/s00253-011-3678-4

A family GH51 α-L-arabinofuranosidase from *Pleurotus ostreatus*: identification, recombinant expression and characterization

Antonella Amore · Angela Amoresano · Leila Birolo · Bernard Henrissat · Gabriella Leo · Angelo Palmese · Vincenza Faraco

Received: 13 July 2011 / Revised: 29 September 2011 / Accepted: 27 October 2011

Abstract An α-L-arabinofuranosidase produced by *Pleurotus ostreatus* (PoAbf) during solid state fermentation on tomato pomace was identified and the corresponding gene and cDNA were cloned and sequenced. Molecular analysis showed that the *poabf* gene carries 26 exons interrupted by 25 introns and has an open reading frame encoding a protein of 646 amino acid residues, including a signal peptide of 20 amino acid residues. The amino acid sequence similar to the other α-L-arabinofuranosidases indicated that the enzyme encoded by *poabf* can be classified as a family 51 glycoside hydrolase. Heterologous recombinant expression of PoAbf was carried out in the yeasts *Pichia pastoris* and *Kluyveromyces lactis* achieving the highest production level of the secreted enzyme (180 mg L^{-1}) in the former host. rPoAbf produced in *P. pastoris* was purified and characterized. It is a glycosylated monomer with a molecular weight of 81,500 Da in denaturing conditions. Mass spectral analyses led to the localization of a single *O*-glycosylation site at the level of Ser160. The enzyme is highly specific for α-L-arabinofuranosyl linkages and when assayed with *p*-nitrophenyl α-L-arabinofuranoside it follows Michaelis–Menten kinetics with a K_M of 0.64 mM and a k_{cat} of 3,010 min^{-1}. The optimum pH is 5 and the optimal temperature 40°C. It is worth noting that the enzyme shows a very high stability in a broad range of pH. The more durable activity showed by rPoAbf in comparison to the other α-L-arabinofuranosidases enhances its potential for biotechnological applications and increases interest in elucidating the molecular bases of its peculiar properties.

Keywords α-L-arabinofuranosidase · Fungus · Recombinant expression · *Pichia pastoris*

A. Amore · A. Amoresano · L. Birolo · G. Leo · A. Palmese · V. Faraco (✉)
Department of Organic Chemistry and Biochemistry,
University of Naples "Federico II",
Complesso Universitario Monte S. Angelo,
via Cintia 4,
80126 Naples, Italy
e-mail: vfaraco@unina.it

A. Amoresano · L. Birolo · V. Faraco
School of Biotechnological Sciences,
University of Naples "Federico II",
Naples, Italy

B. Henrissat
Architecture et Fonction des Macromolécules Biologiques,
UMR6098, CNRS and Universités d'Aix-Marseille I and II,
163 Avenue de Luminy, cedex 9,
13288 Marseille, France

Present Address:
G. Leo
Porto Conte Ricerche Srl,
Loc. Tramariglio, SP 55 Porto Conte/Capo Caccia Km 8.400,
07041 Alghero, Sassari, Italy

Introduction

The α-L-arabinofuranosidases (EC 3.2.1.55) are enzymes that degrade hemicelluloses, such as arabinoxylan, arabinogalactan and L-arabinan, cleaving α-L-arabinofuranosidic linkages and acting synergistically with other enzymes to allow the complete hydrolysis of hemicelluloses and pectins. The α-L-arabinofuranosidases can be found in plant, bacteria and fungi.

There is a growing interest into α-L-arabinofuranosidases due to their application in various industrial processes, such as the production of L-arabinose to be used as food additive, increase of wine aroma by hydrolysis of monoterpenols,

Published online: 13 November 2011

clarification of fruit juices, improvement of the overall quality of bread by delaying the bread staling, improvement of animal feed digestibility, and enhancement of pulp delignification (Numan and Bhosle 2006).

The α-L-arabinofuranosidases are glycoside hydrolases that cleave terminal non-reducing α-L-1,2-, α-L-1,3- and/or α-L-1,5-arabinofuranosyl residues from polymers that contain L-arabinose. Some act on short oligosaccharides, while others are capable of hydrolyzing side-chain arabinose residues from polymeric substrates. α-L-arabinofuranosidases are classified into families GH3, GH43, GH51, GH54 and GH62 of the classification of glycoside hydrolases based on amino acid sequence similarity reported in the CAZy database (http://www.cazy.org/) (Cantarel et al. 2009).

In many cases, microorganisms that utilize hemicelluloses possess α-L-arabinofuranosidases with various substrate specificities and biochemical properties. Some fungal species produce several α-L-arabinofuranosidase isoenzymes.

The white-rot fungus *Pleurotus ostreatus* has been widely studied for its ligninolytic activity due to the expression of a laccase multigene family (Pezzella et al. 2009), while only recently its hemicellulolytic capability has been reported with detection of xylanase activity during solid state fermentation (SSF) on tomato pomace and the observed correlation between production times of xylanase and laccase activities suggested their synergistic action in waste transformation (Iandolo et al. 2011).

In this manuscript an α-L-arabinofuranosidase was identified by proteomic analyses of the enzymes responsible for xylanase activity produced by *P. ostreatus* at the eighth day of SSF on tomato pomace. The gene and cDNA coding for this enzyme named PoAbf were cloned and sequenced. A heterologous expression system was set up for PoAbf in the yeasts *Pichia pastoris* and *Kluyveromyces lactis* achieving the highest production level of the secreted enzyme in the former host. rPoAbf produced in *P. pastoris* was purified and characterized.

Materials and methods

Non-denaturing *polyacrylamide gel electrophoresis*

The zymogram analysis was performed by a modified version of method reported by Ratanakhanokchai et al. (1999). Samples from *P. ostreatus* (Jacq.:Fr.) Kummer (type: Florida; ATCC MYA-2306) suspended in sodium phosphate buffer (50 mM, pH 6.8) were subjected to electrophoresis in an *sodium dodecyl sulfate* (SDS)-10% polyacrylamide gel containing 0.1% birchwood xylan. After electrophoresis, the gel was soaked in 25% (*v*/*v*) isopropanol by gently shaking to remove the SDS and renature the proteins in the gel. The gel was then washed four times for 30 min at 4°C in 0.1 M acetate buffer (pH 5.5). After further incubation for 60 min at 50°C, the gel was soaked in 0.1% Congo red solution for 30 min at room temperature and washed with 1 M NaCl until excess dye was removed from the active band. After, the gel was submerged in 0.5% acetic acid; the background turned dark blue and the activity bands were observed.

In situ mass spectrometry analyses

Slices of interest from the non-denaturing *polyacrylamide gel electrophoresis* (PAGE) on samples from *P. ostreatus* (Jacq.:Fr.) Kummer (type: Florida; ATCC MYA-2306) were destained by washes with 0.1 M NH_4HCO_3 pH 7.5 and acetonitrile, reduced for 45 min with 100 μL of 10 mM dithiothreitol in 0.1 M NH_4HCO_3 buffer pH 7.5 and carboxyamidomethylated for 30 min in the dark with 100 μL of 55 mM iodoacetamide in the same buffer. Tryptic digestion was performed by adding for each slice 100 ng of enzyme in 10 μL of 10 mM NH_4HCO_3 pH 7.5 for 2 h at 4°C. The buffer solution was then removed and 50 μL of 10 mM NH_4HCO_3 pH 7.5 were added and incubated for 18 h at 37°C. Peptides were extracted by washing the gel slices with 10 mM NH_4HCO_3 and 1% formic acid in 50% acetonitrile at room temperature.

The peptide mixtures were filtered by using 0.22 μm PVDF membrane (Millipore) and analysed using a 6520 Accurate-Mass Q-TOF LC/MS System (Agilent Technologies, Palo Alto, CA, USA) equipped with a 1200 HPLC system and a chip cube (Agilent Technologies). After loading, the peptide mixture was first concentrated and washed on 40 nl enrichment column (Agilent Technologies chip), with 0.1% formic acid in 2% acetonitrile as the eluent. The sample was then fractionated on a C18 reverse-phase capillary column (Agilent Technologies chip) at flow rate of 400 nL/min, with a linear gradient of eluent B (0.1% formic acid in 95% acetonitrile) in A (0.1% formic acid in 2% acetonitrile) from 7% to 80% in 50 min. Peptide analysis was performed using data-dependent acquisition of one MS scan (mass range from 300 to 1,800 *m*/*z*) followed by MS/MS scans of the five most abundant ions in each MS scan. MS/MS spectra were measured automatically when the MS signal surpassed the threshold of 50,000 counts. Double and triple charged ions were preferably isolated and fragmented over single charged ions. The acquired MS/MS spectra were transformed in *mz.data* format and used for protein identification with a licensed version of MASCOT 2.1 (Matrix Science, Boston, MA, USA). Raw data from nanoLC–MS/MS analyses were used to query the *P. ostreatus* database (http://genome.jgi-psf.org/PleosPC15_2/PleosPC15_2.home.html; http://genome.jgi-

psf.org/PleosPC9_1/PleosPC9_1.home.html) and the Mascot search parameters were: trypsin as enzyme; allowed number of missed cleavage, 3; carbamidomethyl as fixed modification; oxidation of methionine; pyro-Glu N-term Q as variable modifications; 10 ppm MS tolerance and 0.6 Da MS/MS tolerance; peptide charge from +2 to +3. Spectra with a MASCOT score of <25 having low quality were rejected. The score used to evaluate quality of matches for MS/MS data was higher than 30. Trypsin, dithiothreitol, iodoacetamide, and NH_4HCO_3 were purchased from Sigma. Trifluoroacetic acid-HPLC grade was from Carlo Erba (Milan, Italy). All other reagents and solvents were of the highest purity available from J. T. Baker (Phillipsburg, NJ, USA).

Isolation and sequencing of gene and cDNA

Chromosomal high-molecular weight DNA from *P. ostreatus* was prepared as described by Raeder and Broda (1988). Total RNAs were extracted from lyophilized mycelia by using QIAGEN RNeasy Plant (QIAGEN, Italy) and following manufacturer's instructions. Reverse transcription reaction was performed using MultiScribe™ Reverse Transcriptase (Applied Biosystems, Branchburg, NJ, USA) and the oligonucleotide dT-NotI as primer. Products of the PCR experiments, performed using the gene-specific oligonucleotides (Table 1), were cloned into the p-GEM-T Easy Vector (Promega, Italy). Sequencing by dideoxy chain termination method was performed by PRIMM Sequencing Service (Naples, Italy) using universal and specific oligonucleotide primers.

Analysis of gene and protein sequences

Sequence of gene and cDNA were deposited with the EMBL Data Library under accession numbers HE565355 and HE565356, respectively. Alignments of DNA and of deduced amino acid sequences were generated using ClustalW2 (http://www.ebi.ac.uk/Tools/clustalw2/index.html). Signal peptide prediction was achieved using SignalP V2.0 (http://www.cbs.dtu.dk/services/). Potential *N*-glycosylation sites (Asn-XXX-Ser/Thr) were computed on the NetNGlyc 1.0 server (http://www.cbs.dtu.dk/services/NetNGlyc/), and *O*-GalNAc (mucin type) glycosylation sites were predicted on the OGPET 1.0 server (http://ogpet.utep.edu/OGPET/).

Heterologous expression

Strains, media and plasmids

The *Escherichia coli* strain Top 10 (F-mcrA D (mrrhsdRMS-mcrBC) f80lacZDM15 DlacX74 deoR recA1 araD139 D (ara-leu) 7697 galU galK rpsL (StrR) endA1 nupG) was used in all DNA manipulations. *E. coli* was grown in Luria–Bertani (LB) medium (in grams per liter: 10 bacto tryptone, 10 NaCl, and 5 yeast extract) supplemented, when required, with 100 μg ml^{-1} of ampicillin.

The *K. lactis* strain used for heterologous expression was CMK5 (a *thr lys pgi1 adh3 adh1URA3 adh2URA3*) (Saliola et al. 1999). Plasmid pYG132 (Saliola et al. 1999) was engineered from pKD1, a plasmid originally isolated from *Kluyveromyces drosophilarum* (Falcone et al. 1986) that can replicate stably in *K. lactis*. Both *K. lactis* strain and plasmid were kindly given by Prof. C. Falcone (University "La Sapienza", Rome, Italy). Insert expression is controlled by the ethanol-inducible *KlADH4* promoter (Mazzoni et al. 1992) and the *Saccharomyces cerevisiae* phosphoglycerate kinase (*PGK*) terminator. *K. lactis* was grown in different media, supplemented, when required, with 100 μg ml^{-1} of geneticin G418: YPPG (in grams per liter: 10 yeast extract, 40 bactotryptone and 20 galactose), YPPD (in grams per liter: 10 yeast extract, 40 bactotryptone and 20 Glucose), YPPA (in grams per liter: 10 yeast extract, 40 bactotryptone,

Table 1 List of primers used in the amplification experiments

Primer	Nucleotide sequences	Annealing temperature (°C)
5′ arabin-FW1	ATGTATCGACTCGCTGCACT	60°C
5′ arabin-FW2	ATGTAGGCCTTGGCCCTTG	60°C
5′ arabin-FW3	ATGTTTGAGGTAATGACAATGC	60°C
5′ arabin-FW4	ATGCACGACTCTCCTAGTATCAC	62°C
5′ arabin-FW5	ATGTTCATTCTGAATGATATGCA	60°C
5′ arabin-REV	GCAGCCCCATCGATCGTCAC	66°C
Intarabin-FW	GACGATCGATGGGGCTGC	62°C
Intarabin-REV	GGCGACCGTTTGCAGGCG	62°C
3′ arabin-FW	GCCACGCAATGGAATCACC	60°C
3′ arabin-REV	CTACGACGCCTTGATAGTGA	60°C
contr pYG132 FW	GGGAGAGCTCCAGCAGGGC	66°C
contr pYG132 REV	GGTTCTACCAGCAGGGACGG	66°C

and 20 L-arabinose), YPPX (in grams per liter: 10 yeast extract, 40 bactotryptone and 20 xylose) and YPPWS (in grams per liter: 10 yeast extract, 40 bactotryptone and 20 wheat straw).

The "Easy Select Pichia Expression kit" (Invitrogen, Carlsbad, CA, USA) was used for expression in *P. pastoris* strain X33 following the manufacturer's instructions. pPICZB was chosen as expression vector. Its expression cassette includes the *AOX1* promoter that allows methanol-inducible high level expression and the *AOX1* transcriptional terminator. *P. pastoris* transformants were grown in BMGY (1% yeast extract, 2% peptone, 1%(*v*/*v*) glycerol, 0.00004% biotin, 1.34% yeast nitrogen base with ammonium sulphate and 0.1 M potassium phosphate, pH 6.0) and BMMY, the same as BMGY except that it contains 0.5% (*v*/*v*) methanol instead of glycerol.

Vector construction

The vectors pYG-abf and pPICZ-abf used for recombinant expression of PoAbf in *K. lactis* and *P. pastoris*, respectively were prepared cloning *poabf* cDNA in pYG132 (in *Hin*dIII) under control of the ethanol-inducible *KlADH4* promoter (Piscitelli et al. 2005) and in pPICZB (in *Eco*RI/*Xba*I) under control of methanol-inducible *AOX1* promoter.

Yeasts transformation, cultivation and enzyme production

K. lactis transformation was performed by electroporation with a Bio-Rad MicroPulser apparatus, as specified by the manufacturer. The cells were spread on a YPPG medium containing 100 μg ml^{-1} of geneticin G418, after an overnight incubation at 28°C in YPPG. Agar plate assays on YPPG supplemented with 100 μg ml^{-1} of geneticin G418, 0.5% ethanol and 20 μg ml^{-1} of the chromogenic substrate 4-methylumbelliferyl-a-L-arabinofuranoside (MUA) (Carbosynth, Compton, Berkshire, UK) were used to select transformants. The Abf-expressing clones were identified under UV light by the presence of fluorescence around the colonies after incubation at 28°C for 3 days. 100 uL of ethanol were added each day to the lid of the plate to compensate for its evaporation. The transformed clones were screened for PoAbf production by growing them in 10 mL of selective medium (YPPG containing geneticin) supplemented with 0.5% ethanol, at 28°C on a rotary shaker (150 rpm). Ethanol (0.5%) was added daily and samples were taken at intervals to measure optical density and ABF activity. The best producing clone was chosen to study PoAbf production. Inocula grown on selective YPPG medium were used to seed flasks (250 ml) containing 50 ml of selective medium (YPPG, YPPD, YPPA, YPPX or YPPWS) starting from 0.6 OD600. Cultures were grown at 28°C or 20°C on a rotary shaker for 7 days.

The recombinant plasmid pPICZB–abf was linearized by *Pme*I hydrolysis and introduced into *P. pastoris* strain X33 by electroporation. The electroporation mixture was plated onto YPD–zeocin agar medium. After 72 h at 28°C, the transformant colonies were screened for activity production on plates containing the chromogenic substrate MUA. *P. pastoris* transformants were grown at 30°C in 20 mL of BMGY in an incubator shaker until cell density (OD600) reached the value of 20. The cells were inoculated in 50 mL BMMY in a 250-mL flask to an OD600=1 to start the induction. The culture was kept in an incubator–shaker at 28°C for 15 days (200 rpm) with the addition of 0.25 mL of methanol once a day to maintain induction. For optimizing production, the best producing clone, after growth at 28°C in 20 mL BMGY to an OD600=20, was inoculated in 50 mL of two modified versions of BMMY containing 1% or 1.5% (v/v) methanol and incubated at 28°C or in BMMY at 20°C.

Enzyme activity assays

Xylanase activity assay

Xylanase activity assay was performed according to Bailey et al. (1992). The reaction mixture consisting of 1.8 mL of a 1.0% (*w*/*v*) suspension of birchwood xylan in 50 mM Na citrate at pH 5.3 and 0.2 mL of enzyme dilution (in 50 mM sodium citrate at pH 5.3) was incubated at 50°C for 5 min. Released reducing sugars were determined by DNS method, by adding 3 ml of DNS solution and then incubating the mixture at 95°C for 5 min. Absorbance was measured at 540 nm. One unit of enzyme is defined as the amount of enzyme catalyzing the release of 1 μmol of xylose equivalent per minute.

α-L-arabinofuranosidase activity assay

α-L-Arabinofuranosidase activity was determined according to Yanay and Sato (2000). The activity was measured by spectrophotometric method with *p*-nitrophenyl α-L-arabinofuranoside (pNPA) (Gold Biotechnology, St Louis, MO, USA) as substrate. The assay mixture contained 200 μL of the substrate solution (2 mM pNPA in 50 mM sodium phosphate buffer pH 6.5) and 200 μL of appropriately diluted enzyme solution. After incubation at 30°C for 10 min, the reaction was stopped by the addition of 1.6 mL of 1 M Na_2CO_3. The liberated *p*-nitrophenol in the mixture was measured by spectrophotometry at 405 nm (ε=17,000 M^{-1} cm^{-1}). One unit of α-L-arabinofuranosidase activity was defined as the amount of enzyme releasing 1 μmol of *p*-nitrophenol per minute in the reaction mixture under these assay conditions.

PoAbf purification

Proteins secreted by rPoAbf expressing *P. pastoris* were precipitated from the medium (after cell removal by centrifugation at 5,000×*g* for 10 min) by the addition of $(NH_4)_2SO_4$ up to 80% saturation at 4°C and centrifugation at 10,000×*g* for 30 min. The precipitate was resuspended in 0.02 M Tris–HCl pH 7.5 and loaded on HiTrap Phenyl FF high sub (GE Healthcare, Uppsala, Sweden) equilibrated in buffer A (0.02 M Tris–HCl, 1.2 M $(NH_4)_2SO_4$, pH 7.5), and the proteins were eluted isocratically with buffer B (0.02 M Tris–HCl pH 7.5). Fractions containing activity were combined and concentrated on an Amicon PM-10 membrane and analyzed by SDS-PAGE.

Protein concentration determination

Protein concentration was determined by the method of Lowry et al. (1951), using the BioRad Protein Assay (BioRad Laboratories S.r.l., Segrate, MI, Italy), with bovine serum albumin as standard.

Determination of molecular mass

The molecular mass of rPoAbf produced by *P. pastoris* was determined by using a Superdex 200 HR 10/30 column (GE Healthcare, Uppsala, Sweden). The column was eluted with 50 mM Tris–HCl buffer, pH 7.0, containing 150 mM NaCl. The calibration was performed with thyroglobulin (669 kDa), apoferritin (440 kDa) alcohol dehydrogenase (150 kDa), bovine serum albumin (67 kDa), and ovalbumin (45 kDa) as standards.

Optimum temperature and temperature stability

To determine the optimum temperature of the purified enzyme, the substrate of the activity assay (2 mM pNPA) was dissolved in 50 mM sodium phosphate buffer pH 6.5 and the incubation (10 min) was performed at 30°C, 40°C, 50°C, 55 C and 60°C. The temperature stability of rPoAbf was studied by incubating the purified enzyme preparation in 50 mM citrate phosphate buffer pH 5.0, at 30°C, 40°C, 50°C and 60°C. The samples withdrawn were assayed for residual α-arabinosidase activity performing incubation (10 min) at 40°C. The experiments were performed in duplicate and reported values are representative of three experiments.

Optimum pH and pH stability

To determine the optimum pH of the purified enzyme, the substrate of the activity assay (2 mM pNPA) was dissolved in citrate phosphate buffers (McIlvaine 1921) with pH values between 3.0 and 7.0 and in 50 mM Tris–HCl with pH values 7.0 and 8.0 and the incubation (10 min) was performed at 40°C. The pH stability of the purified enzyme preparation was studied by diluting it in citrate phosphate buffers, pH 3–8 and incubating at 25°C. From time to time, samples were withdrawn and immediately assayed for residual α-arabinosidase activity performing incubation (10 min) at 40°C. The experiments were performed in duplicate and reported values are representative of three experiments.

Assays of enzyme specificity

Activity of PoAbf was assayed against the substrates *p*NP-α-L-arabinofuranoside (Carbosynth, Compton, Berkshire, UK), *p*NP-β-D-xylopyranoside (Carbosynth, Compton, Berkshire, UK), *p*NP-α-D-glucopyranoside (Carbosynth, Compton, Berkshire, UK) *p*NP-β-D-glucopyranoside (Carbosynth, Compton, Berkshire, UK) and *o*NP-β-D-galactopyranoside (Carbosynth, Compton, Berkshire, UK) at concentrations in the range 0.1–6 mM in citrate phosphate buffer (McIlvaine, 1921) at pH 5.

Activity against the natural substrates CM-linear arabinan and larch arabinogalactan, and the arabinooligosaccharides 1,5-α-arabinotriose and 1,5-α-arabinohexaose (Megazyme International Ireland, Co. Wicklow, Ireland) was assayed measuring the liberation of arabinose as equivalent of galactose, by means of the D-galactose/lactose kit (Megazyme International Ireland, Co. Wicklow, Ireland), following the manufacturer's instructions. 600 mU (measured on pNPA) of PoAbf were incubated with 0.2% arabinans in 100 mM sodium acetate buffer pH 4.6, at 37°C for 72 h (final volume, 300 μL). Arabinooligosaccharides were dissolved in 100 mM sodium acetate buffer pH 4.6 and incubated with 7 mU (measured on pNPA) of PoAbf at 37°C for 1 h (final volume 300 μL). The activity of PoAbf against AZO-wheat arabinoxylan (Megazyme International Ireland, Co. Wicklow, Ireland) was assayed following supplier's instructions. The experiments were performed in duplicate and reported values are representative of three experiments.

Determination of k_{cat} and K_M

For determination of the Michaelis–Menten constants K_M and k_{cat}, the activity assay was performed at pNPA concentrations from 0.1 mM to 2.0 mM at pH 5.0 performing incubation of 10 min at 40°C. The experiments were performed in duplicate and reported values are representative of three experiments.

Mass spectrometry analyses of rPoAbf

α-L-arabinofuranosidase was dissolved in denaturant buffer (Tris 300 mM pH 8.0, urea 6 M, EDTA 10 mM), and disulfide bridges were reduced with DTT (tenfold molar

excess on the Cys residues) at 37°C for 2 h and then alkylated by adding IAM (fivefold molar excess on thiol residues) at room temperature for 30 min in the dark. Protein sample was desalted by size exclusion chromatography on a Sephadex G-25M column (GE Healthcare). Fractions containing protein were lyophilized. Lyophilized fractions were dissolved in 10 mM AMBIC buffer pH 8.0. Enzymatic digestion was performed using both trypsin and V8 protease using an enzyme/substrate ratio of 1/50 (*w*/*w*) at 37°C for 16 h. MALDI–MS experiments were performed on a Voyager-DE STR MALDITOF mass spectrometer (Applied Biosystems, Framingham, MA, USA) equipped with a nitrogen laser (337 nm). Of the total mixture, 1 μL was mixed (1/1, *v*/*v*) with a 10-mg mL^{-1} solution of α-cyano-4-hydroxycinnamic acid in ACN/50 mM citrate buffer, 70/30 (*v*/*v*). Spectra were acquired using a mass (*m*/*z*) range of 400–5,000 amu.

Results

Enzyme identification

The native PAGE showed that two different xylanase bands can be detected analyzing the second and the eighth days of *P. ostreatus* (Jacq.:Fr.) Kummer (type: Florida; ATCC MYA-2306)-SSF on tomato processing residues, when maximum level of xylanase activity—related to any enzyme acting on any part of xylan—was reached (Iandolo et al. 2011). The activity positive protein band with lower electrophoretic mobility produced at the eighth day was excised from a preparative, non-denaturing gel and analyzed by proteomics. By comparison with genome sequences of the monokaryon strains *P. ostreatus* PC15 (http://genome.jgi-psf.org/PleosPC15_2/PleosPC15_2.home.html) and PC9 (http://genome.jgi-psf.org/PleosPC9_1/PleosPC9_1.home.html), the protein was identified as an α-L-arabinofuranosidase—having the ID 29657 in *P. ostreatus* PC15 and 94793 in *P. ostreatus* PC9 strains—that was named PoAbf. The fact that, even if xylanase activity had been found, an arabinofuranosidase was identified can be explained considering that arabinoxylan is used as a substrate and that this substrate contains arabinose side branches on which the enzyme seems to act. The time course of α-L-arabinofuranosidase activity was then analyzed during *P. ostreatus* SSF on tomato processing waste by using the pNPA as substrate, thus demonstrating the presence of a peak of around 0.18 Ug^{-1} dry matter at the sixth day of fermentation.

Sequences of the *poabf* gene and cDNA and its derived protein

Analysis of genome sequence of the monokaryon *P. ostreatus* strain PC15 showed that the gene with ID 29657 is interrupted at the 5′ region and does not contain any start codon. On the other hand, a complete coding sequence was identified in the monokaryon *P. ostreatus* strain PC9 for the gene with ID 94793. However, analysis of the deduced amino acid sequence corresponding to the latter gene according to the annotated version showed that it lacks the signal peptide, in spite of the demonstrated extracellular nature of the analyzed α-L-arabinofuranosidase. The analysis of genome sequence of PC9 strain upstream of the annotated start codon allowed design of different oligonucleotides to be used in PCR experiments to amplify cDNA coding for the investigated protein in the dikaryon strain *P. ostreatus* ATCC MYA-2306. This cDNA was synthesized by PCR on RNA extracted from *P. ostreatus* grown on tomato processing residues at the eighth day of SSF using as primers oligonucleotides (Table 1) whose sequences were designed on the basis of genome sequence to amplify three overlapped fragments. By carrying out PCR experiments of the 5′ region with different primer oligonucleotides at the 5′ end (5′ arabin-FW1-5) in combination with the same oligonucleotide at the 3′ end (5′ arabin-REV), it was possible to ascertain the start codon of *poabf*, considering which oligonucleotide gave the most 5′ extended region. *poabf* gene was also synthesized by PCR on *P. ostreatus* ATCC MYA-2306 genome.

The cDNA coding for PoAbf has an open reading frame of 1,941 bp and it codes for a protein of 646 amino acids. A unique signal peptide sequence of 20 amino acids could be singled out, which satisfied criteria for a "minimal" signal sequence (von Heijne 1985).

The mature protein is 626 amino acids in length and has a calculated molecular mass of 68,912.51 Da. The coding sequence is interrupted by 25 introns as detected when the cDNA and gene sequences were compared. The intron size ranges from 46 to 112 bp. All introns are flanked by the consensus sequences GT.AG (Kupfer et al. 2004). No potential *N*-glycosylation site (Asn-X-Ser) was found in PoAbf deduced amino acid sequence, while Ser160 and Thr187 were identified as potential *O*-GalNAc (mucin type) glycosylation sites within the patterns T-n-G**S**T-n-VS and T-n-RTT-n-PS, respectively.

Similarity searches were performed with Basic Local Alignment Search Tool (BLAST) using the deduced amino acid sequence of PoAbf as the query and a library made with the sequences of experimentally characterized family GH51 enzymes (http://www.cazy.org/GH51_characterized.html) extracted from the Carbohydrate-active enzymes database (http://www.cazy.org/, Cantarel et al. 2009). The best hit (62% identities) was with an α-L-arabinofuranosidase from *Meripilus giganteus* (EMBL accession CAL81200.1; Sørensen et al. 2006), followed by several enzymes from *Aspergillus* spp. (EMBL accessions BAB96815.1, BAB21568.2, AAC41644.1, etc.) displaying ~35% identities

with PoAbf. Plant and bacterial GH51 ABFs gave lower scores, typically less than 30% identities. The distribution of genes encoding enzymes similar to PoAbf was examined by performing a BLAST search (conducted on Sept 22, 2011) on the non-redundant protein sequence databank of the NCBI. The results revealed that only a limited number of fungi have GH51 enzymes with over 50% identity to PoAbf, namely *Leucoagaricus gongylophorus*, *Serpula lacrymans*, *Schizophyllum commune*, *Postia placenta* and *Coprinopsis cinerea*. All other fungal hits showed less than 40% sequence identity with PoAbf, hinting at possible differences in substrate or product specificities.

The family GH51 glycoside hydrolases have a catalytic mechanism with retention of configuration with two glutamates involved in catalysis (one acting as an acid/base and the other as a nucleophile) (Henrissat et al. 1995). Three-dimensional structures of family GH51 ABFs have been obtained for two bacterial enzymes, namely from *Geobacillus stearothermophilus* (Hövel et al. 2003) and *Clostridium thermocellum* (Taylor et al. 2006). The catalytic domain is a (β/α)8 barrel, and the active site glutamates are found at the C-termini of strands β-4 (acid/base) and β-7 (nucleophile) (Taylor et al. 2006). Based on this information and the sequence alignment, the proposed catalytic glutamates of PoAbf are located at positions 364 (acid/base) and 471 (nucleophile).

Recombinant expression system

The enzyme α-L-arabinofuranosidase secreted by *P. ostreatus*, PoAbf, was expressed in *K. lactis* and *P. pastoris*. The *poabf* cDNA sequence was optimized for yeast recombinant expression by adapting it to the yeast codon usage and avoiding its hydrolysis by the endonucleases adopted for the cloning. Synthesis of this cDNA was performed by MrGene (http://mrgene.com).

K. lactis transformants

When the production of recombinant α-L-arabinofuranosidase in *K. lactis* was assayed on plates containing MUA as substrate, around 20% of transformants caused a fluorescence emission after 3 days. Eighteen recombinant clones were grown in liquid cultures and evaluated for Abf production by assay with pNPA. The amount of extracellular α-L-arabinofuranosidase activity produced after 7 days ranged from (0 to 3.50)$\times 10^{-3}$U mL^{-1}. The best producing clone was used to optimize production of rPoAbf and study the time course of recombinant α-L-arabinofuranosidase. Cultures were carried out to improve α-L-arabinofuranosidase production, varying the growth conditions. Decreasing the growth temperature from 28°C to 20°C led to a 10-fold increase of activity production. Comparison of growth time courses at the two temperatures showed only a slight delay of yeast growth due to the temperature decrease. Glucose, galactose, L-arabinose, xylose and wheat straw were tested as carbon sources. The maximum activity level was obtained at the seventh day with all the tested substrates. The highest production of secreted activity was achieved at 20°C in the medium containing glucose (62.0×10^{-3}U mL^{-1}), followed by those containing xylose (49.72×10^{-3}U mL^{-1}), galactose and wheat straw (35.55×10^{-3}U ml^{-1}) and arabinose (20.36×10^{-3}U mL^{-1}).

P. pastoris transformants

All the *P. pastoris* transformant clones were proven to secrete recombinant α-L-arabinofuranosidase activity as

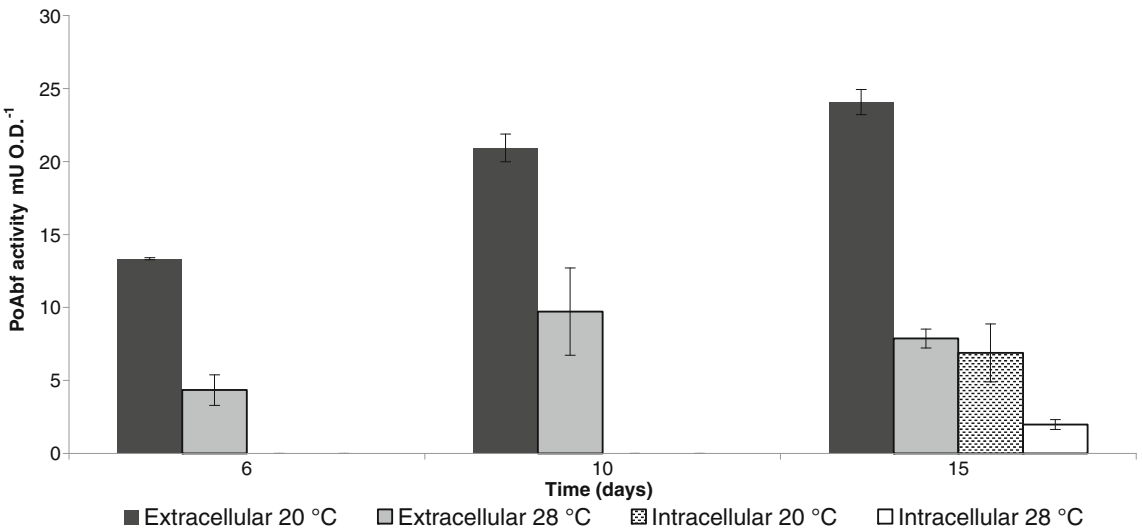


Fig. 1 Effect of growth temperature on α-arabinofuranosidase activity levels produced by *P. pastoris* recombinant clone

shown by assay on plates containing MUA as substrate, giving fluorescence emission after 3 days. Seventeen recombinant clones were grown in liquid cultures and evaluated for Abf production. The maximum amount of α-L-arabinofuranosidase activity produced after 15 days was of 30×10^{-2}U mL^{-1}, corresponding to 8.3 mU $O.D.^{-1}$, in culture supernatant and of 1.2 mU $O.D.^{-1}$ in the cellular extracts.

The best producing clone was used to optimize rPoAbf production, study the time course of α-L-arabinofuranosidase production and characterize the recombinant enzyme. It was shown that rPoAbf production increases linearly for at least 1 month. Cultures were carried out to improve α-L-arabinofuranosidase production, varying the growth conditions. Increasing methanol concentration from 0.5 to 1 and

Table 2 MALDI–MS analysis of the reduced and carboxamidomethylated rPoAbf digested with trypsin

Peptide	*m/z*	Sequence
1-3	469.28	MYR
352-357	643.40	RAALGR
104-110	747.43	FTVPAGR
255-261	758.46	GQTVPTR
543-551	1033.60	YIPSTLPSR
481-489	1041.53	AAYMIGMER
383-391	1073.65	WPAFVNALR
552-561	1095.60	TGTAFWGVTR
244-254	1297.68	FFRFPGGNNLE
91-103	1369.79	SVPVSNSLPNSLR
226-239	1447.74	ANGMRDDIATALAE
156-170	1546.90	TTSG**S**TIVSQTVPIR
111-125	1547.76	TGQVGFANSGFFGMK
247-261	1586.81	FPGGNNLEGQTVPTR
171-186	1717.95	GASTSWQQISVSLTPR
358-372	1723.92	AQPFKLTYVEIGNED
628-644	1786.95	TFTFAAAGFSVNVITIK
240-254	1782.00	MKPKFFRFPGGNNLE
226-243	1931.99	ANGMRDDIATALAEMKPK
44-61	1945.01	DISHSGDGGLYAELLQNR
450-468	1965.03	YAAISTNPNDIFGSPANGR
534-551	2027.07	LFSTNMGNEYIPSTLPSR
244-261	2037.06	FFRFPGGNNLEGQTVPTR
552-571	2136.10	TGTAFWGVTRNTATGEIIIK
448-468	2151.11	GEYAAISTNPNDIFGSPANGR
441-468	2961.42	NGVTFYEGEYAAISTNPNDIFGSPANGR
469-489	2281.07	LTFSTMQSASGEAAYMIGMER
104-125 + Oxidized Met	2292.19	FTVPAGRTGQVGFANSGFFGMK
134-155	2345.26	GSFFVKFPTASSFTGNLIVGLR
363-382	2404.19	LTYVEIGNEDFFASDTYTYR
372-398	3167.65	DFFASDTYTYRWPAFVNALRGAFPNLK
140-170 + GalNacHex	3572.62	FPTASSFTGNLIVGLRTTSG**S**TIVSQTVPIR
140-170 + $GalNacHex_2$	3734.80	FPTASSFTGNLIVGLRTTSG**S**TIVSQTVPIR
490-525	3865.11	NSDIVFAASYAPLLGHVAGSQWTPNLIAFD TGNVYR
187-223	3967.06	TTPPSTDNQFFVTIDGAAASGQTINFAMFS LFPPTFK
51-90	4111.13	GGLYAELLQNRALQQVTPNTAASLNAWSAI NGGQISVVAD
62-103	4247.44	ALQQVTPNTAASLNAWSAINGGQISVVADS VPVSNSLPNSLR
126-170	4722.30	IVASSTYKGSFFVKFPTASSFTGNLIVGLR TTSGSTIVSQTVPIR
441-489	5224.16	NGVTFYEGEYAAISTNPNDIFGSPANGRLT FSTMQSASGEAAYMIGMER

Ser160 in bold is glycosylated

1.5% causes a decrease of rPoAbf production by 2.5- and 4-fold, respectively, while decreasing the growth temperature from 28°C to 20°C lead to an increase of activity production (up to 1.1 UmL^{-1} after 15 days) (Fig. 1). Enzyme activity was also found in intracellular protein fraction at both analyzed temperatures (Fig. 1).

rPoAbf characterization

The enzyme purified to apparent homogeneity was subjected to structural and functional characterization as described below. The estimated molecular weight deduced from SDS-PAGE is 81,500 Da, considerably larger than that deduced from cDNA sequence (data not shown), probably due to an anomalous hydrodynamic behavior during electrophoretic separation on SDS-PAGE as expected from a glycosylated protein (see below). Using chromatography in a calibrated Superdex 200 HR 10/30 column (GE Healthcare, Uppsala, Sweden), a molecular weight of 39,000 Da was obtained for the native enzyme (data not shown). These results indicate that PoAbf is a monomer. The reason for the smaller apparent molecular weight in gel filtration could be retarded migration of the α-arabinosidase in the gel matrix, as was also reported for the α-arabinosidases and xylanases of *Trichoderma reesei* (Lappalainen 1986; Poutanen 1988). The presence of a galactose-binding domain like in the enzyme could be the reason for its binding to the column material.

Mass spectrometry analyses of rPoAbf allowed us to map 77.9% of the total amino acid sequence and the results are summarized in Table 2. Mass spectral analyses led to the localization of a single *O*-glycosylation site at level of Ser160, while Thr187 was found unmodified. Figure 2 shows the partial MALDI–MS spectrum obtained from the analysis of the tryptic peptide mixture of rPoAbf. The spectrum shows the occurrence of a signal at *m*/*z* 3572.62 assigned to peptide 140-170 modified with an oligosaccharide chain containing a GalNac and an Hexose moiety; moreover it is evident that the presence of a signal at *m*/*z* 3743.80 was attributed to the same peptide carrying a GalNacHex2 oligosaccharide moiety thus indicating the heterogeneity of *O*-glycosylation at Ser160.

The enzyme follows Michaelis–Menten kinetics towards pNPAra: the K_M for this substrate is 0.64±0.11 mM and the k_{cat} 3,010±145 min^{-1}. The pH optimum for the enzyme (assayed in a range from 3–8) is 5 (Fig. 3a) and the optimal temperature is 40°C (Fig. 3b). PoAbf shows a half life of 17 h at 50°C and few minutes at 60°C, while it retains 100% of its activity for at least 1 week at 30°C and 40°C. PoAbf shows very high stability in a broad range of pH ($T_{1/2}$ pH 3=24 days; $T_{1/2}$ pH 5=51 days; $T_{1/2}$ pH $7_{citrate–phosphate\ buffer}$=38 days; $T_{1/2}$ pH $7_{Tris–HCl\ buffer}$=42 days; $T_{1/2}$ pH 8=38 days).

The hydrolyzing ability of rPoAbf was tested versus a series of other nitrophenyl glycosides: *p*NP-β-D-xylopyranoside, *p*NP-α-D-glucopyranoside, *p*NP-β-D-glucopyranoside, and *o*NP-β-D-galactopyranoside. Among these compounds, only *p*NP-β-D-glucopyranoside was recognized by the

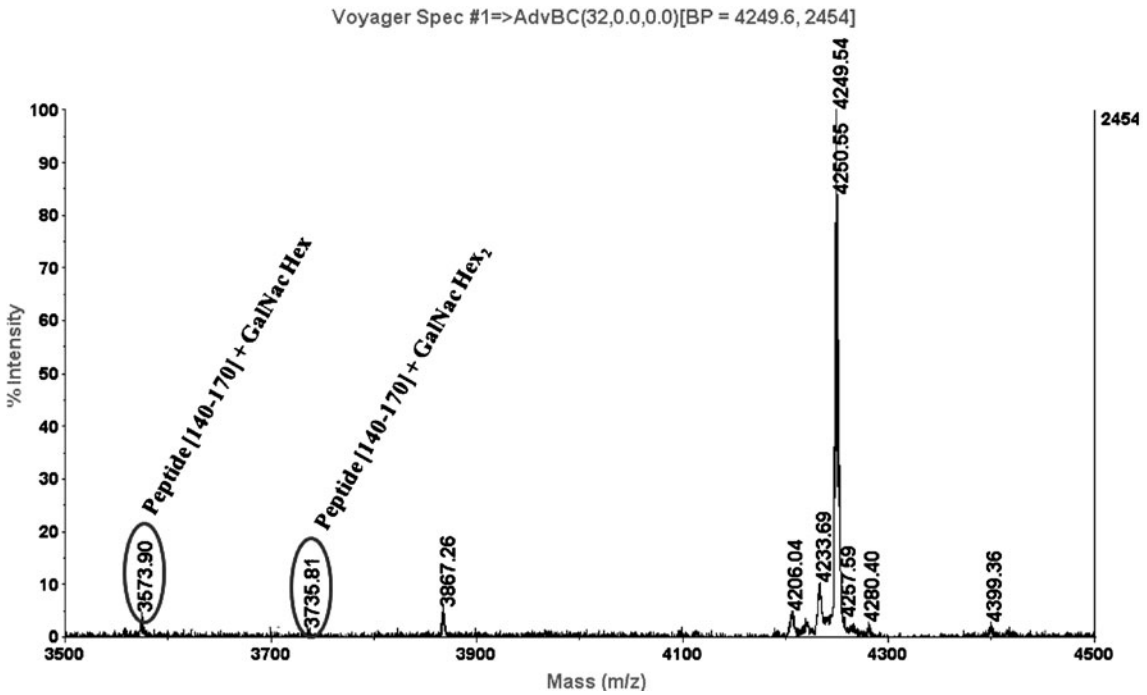


Fig. 2 Partial MALDI–MS spectrum obtained from the analysis of the tryptic peptide mixture of rPoAbf

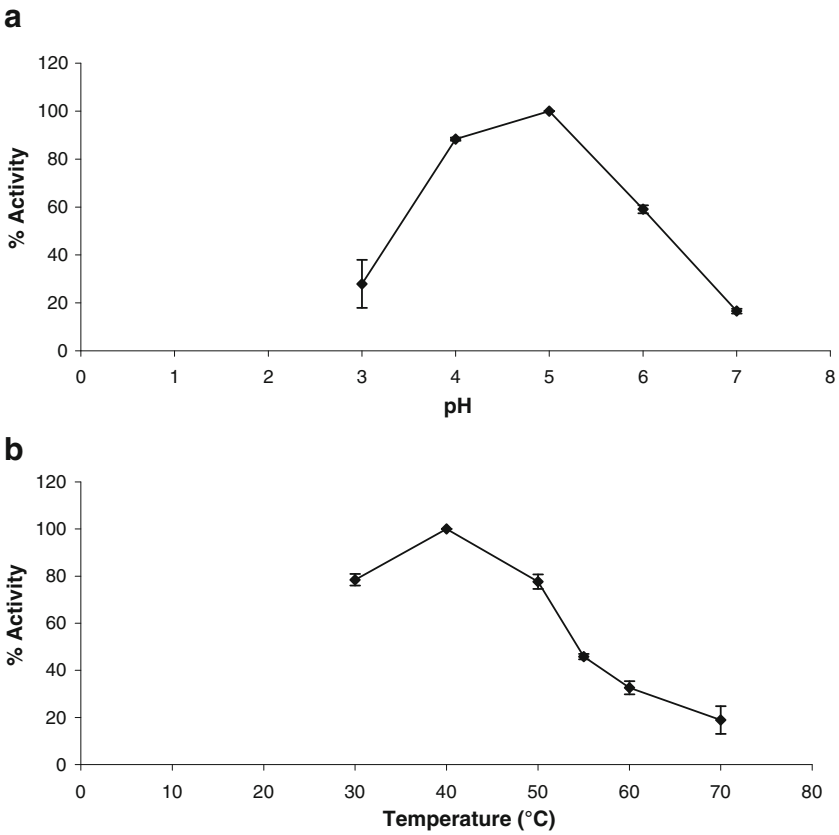


Fig. 3 Effect of pH (**a**) and temperature (**b**) on the activity of rPoAbf towards pNPA as substrate. The percent activity reported in **a** and **b** was calculated as ratio to the maximum activity at optimum pH (5) and temperature (40°C), respectively

enzyme and a K_M of 3.79±0.12 mM and a k_{cat} of 9.3±0.5 min^{-1} were measured for this substrate.

The enzyme was shown able to hydrolyze both the tested arabinooligosaccharides 1,5-α-arabinotriose and 1,5-α-arabinohexaose (Table 3). PoAbf is active on CM-linear arabinan (Table 3), while it does not hydrolyze larch arabinogalactan. Moreover, when the enzyme was incubated with the AZO-wheat arabinoxylan, it was shown to possess an endo-1,4-β-xylanase activity of around 0.55±0.08 UmL^{-1}.

Discussion

To the best of our knowledge, this is the first report on the cloning, recombinant expression, purification and characterization of α-L-arabinofuranosidase from the genus of *Pleurotus*. This α-L-arabinofuranosidase was identified analyzing the enzymes responsible for xylanase activity produced by *P. ostreatus* SSF on tomato pomace. The gene and cDNA coding for this enzyme named PoAbf were cloned and sequenced. Analysis of the amino acid sequence of PoAbf places the enzyme in family 51 of the glycoside hydrolases.

A heterologous expression system was set up in the yeasts *P. pastoris* and *K. lactis* using the leader sequence of PoAbf in both the expression systems and the highest production level of the secreted enzyme (180 mg L^{-1}) was achieved in the former host. Decreasing growth temperature from 28°C to 20°C caused an increase of recombinant expression level of extracellular enzyme of around three-fold. Comparison of production levels of intracellular recombinant enzyme at the analyzed temperatures allowed us to demonstrate that the low temperature favors the production but it does not improve protein secretion. Many studies have demonstrated that temperature is a critical parameter in *P. pastoris* growth and recombinant protein production, in both batch and fed-batch systems and it has been reported that reducing induction temperature to 20°C or even lower values is beneficial for the efficient expression of heterologous proteins in *P. pastoris*. A common opinion accounting for this effect lies in the fact that lowering temperature could at least help to reduce cell

Table 3 Arabinose liberation from natural substrates and arabinooligosaccharides

Substrate	Amount of released sugar (μg/mL)
Larch arabinogalactan	0
CM-linear arabinan	327±24
Arabinotriose	112±18
Arabinohexaose	102±26

Equivalents of galactose were measured as described in section "Material and methods"

skeleton lysis and protease secretion when cultivated on methanol, besides reducing cell death. Jin et al. (2010) reported that lowering induction temperature from 30°C to 20°C causes a 100-fold increase of recombinant production of Porcine interferon-α in *P. pastoris*. A reduction of temperature from 30°C to 25°C or 20°C resulted in a 2- and 3-fold increase of the specific productivity of the 3H6 Fab fragment, respectively (Dragosits et al. 2009). Li et al. (2001) showed that the increase of the recombinant production of herring antifreeze protein at lower temperature might benefit from the better protein folding. In fact, cultivating the cells at lower growth temperatures induces the rate of protein synthesis and thus may allow the nascent peptide chains to fold properly. By lowering the induction temperature to 22°C, also recombinant polygalacturonate lyase activity exhibited a 2.9-fold increase compared to cultivation at 30°C (Wang et al. 2009).

In all the other expression systems so far set up in *P. pastoris*, α-L-arabinofuranosidase was expressed under the control of yeast signal peptide. When compared with the other α-L-arabinofuranosidase recombinant expression systems, the system hereby reported showed similar or better yields, in terms of both units per milliliter and milligrams per liter. Particularly, our system was proven to be more efficient than the other recombinant expression systems so far reported in *P. pastoris*, where an expression yield of 20 mg L^{-1} for the α-L-arabinofuranosidase B from *Aspergillus kawachii* (Miyanaga et al. 2004) and of 0.2 UmL^{-1} for the α-L-arabinofuranosidase from *Trichoderma koningii* (Wan et al. 2007) was achieved. Yields obtained in *S. cerevisiae* expression systems ranged from 0.013 UmL^{-1} for the α-L-arabinofuranosidase from *Aureobasidium pullulans* (Wan et al. 2007) to 1.40 UmL^{-1} for *Aspergillus niger* α-L-arabinofuranosidase (Crous et al. 1996). Similar activity levels were achieved with *E. coli* expression systems, giving 0.94 UmL^{-1} (Lee et al. 2011) for the α-L-arabinofuranosidases from *Penicillium* sp. LYG 0704, 1.48 Umg^{-1} for the α-L-arabinofuranosidase from *Fusarium oxysporum* f. sp. *dianthi* (Chacòn-Martìnez et al. 2004) and 1.2 UmL^{-1} (Shi et al. 2010) for the α-L-arabinofuranosidase from *Streptomyces* sp. S9.

rPoAbf produced in *P. pastoris* was purified and characterized, and it was shown to have some properties similar to the other members of family 51 ABFs (Numan and Bhosle 2006). Their subunit molecular weight generally ranges from 56–81 kDa. All enzymes show an acidic pH optimum and a value of K_M for pNPAra ranging from 0.5–0.8 mM.

PoAbf revealed to be a versatile enzyme able to work on arabinooligosaccharides, with a higher affinity for the shorter ones, and on the natural polysaccharides linear arabinan and arabinoxylan, displaying both exo- and endoxylanase activities. It is worth noting that PoAbf shows very high stability in a broad range of pH, mostly at acidic pH, while most of other α-L-arabinofuranosidases were proven to be less stable, such as recombinant arabinofuranosidase from *Streptomyces sp. 9* that totally lost its activity after incubation for 1 h at pH 4 (Shi et al. 2010), or *A. pullulans* α-L-arabinofuranosidase that shows a half life of 29 min at pH 3 (Wan et al. 2007). PoAbf stability at pH 3 is also higher than stability of α-L-arabinofuranosidase from *T. reesei* (Poutanen 1988). The highly durable activity of rPoAbf enhances its potential for biotechnological applications and increases interest in this enzyme as a candidate for mutagenesis experiments aimed at elucidating the molecular bases of its peculiar properties.

Acknowledgments This work was supported by grant from the Ministero dell'Università e della Ricerca Scientifica (Progetto di ricerca industriale ENERBIOCHEM PON01_01966, finanziato nell'ambito del Programma Operativo Nazionale Ricerca e Competitività 2007–2013 D. D. Prot. n. 01/Ric. del 18.1.2010).

References

Bailey MJ, Biely P, Poutanen K (1992) Interlaboratory testing of methods for assay of xylanase activity. J Biotechnol 23:257–270

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The carbohydrate-active enzymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res 37:233–238

Chacòn-Martìnez CA, Anzola JM, Rojasa A, Hernandez F, Junca H, Ocampo W, Del Portillo P (2004) Identification and characterization of the α-L-arabinofuranosidase B of *Fusarium oxysporum* f. sp. *dianthi*. Phys Mol Plant Pathol 64:201–208

Crous JM, Pretorius IS, van Zyl WH (1996) Cloning and expression of the α-L-arabinofuranosidase gene (*ABF2*) of *Aspergillus niger* in *Saccharomyces cerevisiae*. Appl Microbiol Biotechnol 46:256–260

Dragosits M, Stadlmann J, Albiol J, Baumann K, Maurer M, Gasser B, Sauer M, Altmann F, Ferrer P, Mattanovich D (2009) The effect of temperature on the proteome of recombinant *Pichia pastoris*. J Proteome Res 8:1380–1392

Falcone C, Saliola M, Chen XJ, Frontali L, Fukuhara H (1986) Analysis of a 1.6-micron circular plasmid from the yeast *Kluyveromyces drosophilarum*: structure and molecular dimorphism. Plasmid 15:248–252

Henrissat B, Callebaut I, Fabrega S, Lehn P, Mornon JP, Davies G (1995) Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. Proc Nat Acad Sci U S A 92:7090–7094

Hövel K, Shallom D, Niefind K, Belakhov V, Shoham G, Baasov T, Shoham Y, Schomburg D (2003) Crystal structure and snapshots along the reaction pathway of a family 51 α-L-arabinofuranosidase. EMBO J 22:4922–4932

Iandolo D, Piscitelli A, Sannia G, Faraco V (2011) Enzyme production by solid substrate fermentation of *Pleurotus ostreatus* and *Trametes versicolor* on tomato pomace. Appl Biochem Biotechnol 163(1):40–51

Jin H, Liua G, Yeb X, Duana Z, Li Z, Shia Z (2010) Enhanced porcine interferon-α production by recombinant *Pichia pastoris* with a

combinational control strategy of low induction temperature and high dissolved oxygen concentration. Biochem Eng J 52:91–98

Kupfer DM, Drabenstot SD, Buchanan KL, Lai H, Zhu H, Dyer DW, Roe BA, Murphy JW (2004) Introns and splicing elements of five diverse fungi. Eukaryot Cell 3(5):1088–1100

Lappalainen A (1986) Purification and characterization of xylanolytic enzymes from *Trichoderma reesei*. Biotechnol Appl Biochem 8:437–448

Lee D-S, Wi S-G, Lee Y-G, Cho E-J, Chung B-Y, Bae H-J (2011) Characterization of a new α-L-arabinofuranosidase from *Penicillium* sp. LYG 0704, and their application in lignocelluloses degradation. Mol Biotechnol. doi:10.1007/s12033-011-9396-4

Li Z, Xiong F, Lin Q, d'Anjou M, Daugulis AJ, Yang DSC, Hew CL (2001) Low-temperature increases the yield of biologically active herring antifreeze protein in *Pichia pastoris*. Protein Expr Purif 21:438–445

Lowry OH, Rosebrough NJ, Farr AL, Randall RJ (1951) Protein measurement with the Folin phenol reagent. J Biol Chem 193:265–275

Mazzoni C, Saliola M, Falcone C (1992) Ethanol-induced and glucose-insensitive alcohol dehydrogenase activity in the yeast *Kluyveromyces lactis*. Mol Microbiol 6:2279–2286

McIlvaine TC (1921) A buffer solution for colorimetric comparison. J Biol Chem 49:183–186

Miyanaga A, Koseki T, Matsuzawa H, Wakagi T, Shoun H, Fushinobu S (2004) Expression, purification, crystallization and preliminary X-ray analysis of α-L-arabinofuranosidase B from *Aspergillus kawachii*. Acta Crystallogr D Biol Crystallogr D60:1286–1288

Numan MT, Bhosle NBJ (2006) α-L-arabinofuranosidases: the potential applications in biotechnology. J Ind Microbiol Biotechnol 33(4):247–260

Pezzella C, Autore F, Giardina P, Piscitelli A, Sannia G, Faraco V (2009) The *Pleurotus ostreatus* laccase multi-gene family: isolation and heterologous expression of new family members. Curr Genet 55(1):45–57

Piscitelli A, Giardina P, Mazzoni C, Sannia G (2005) Recombinant expression of *Pleurotus ostreatus* laccases in *Kluyveromyces lactis* and *Saccharomyces cerevisiae*. Appl Microbiol Biotechnol 69:428–439

Poutanen K (1988) An α-L-arabinofuranosidase of *Trichoderma reesei*. J Biotechnol 7:271–282

Raeder V, Broda P (1988) Preparation and characterization of DNA from lignin degrading fungi. Methods Enzymol 161B:211–220

Ratanakhanokchai K, Kyu KL, Tanticharoen M (1999) Purification and properties of a xylan-binding endoxylanase from alkaliphilic *Bacillus* sp. strain K-1. Appl Environ Microbiol 65:694–697

Saliola M, Mazzoni C, Solimando N, Crisà A, Falcone C, Jung G, Fleer R (1999) Use of the *KlADH4* promoter for ethanol-dependent production of recombinant human serum albumin in *Kluyveromyces lactis*. Appl Environ Microbiol 65:53–60

Shi PJ, Li N, Yang PL, Wang Y, Luo HY, Bai Y, Yao B (2010) Gene Cloning, Expression and Characterization of a Family 51 α-L-arabinofuranosidase from *Streptomyces* sp. S9. Appl. Biochem. Biotechnol 162(3):707–18

Sørensen HR, Jørgensen CT, Hansen CH, Jørgensen CI, Pedersen S, Meyer AS (2006) A novel GH43 α-L-arabinofuranosidase from *Humicola insolens*: mode of action and synergy with GH51 α-L-arabinofuranosidases on wheat arabinoxylan. Appl Microbiol Biotechnol 73(4):850–861

Taylor EJ, Smith NL, Turkenburg JP, D'Souza S, Gilbert HJ, Davies GD (2006) Structural insight into the ligand specificity of a thermostable family 51 arabinofuranosidase, Araf51, from *Clostridium thermocellum*. Biochem J 395:31–37

von Heijne G (1985) Signal sequences. The limits of variation. J Mol Biol 184:99–105

Wan C-F, Chen W-H, Chen C-T, Chang MD-T, Lo L-C, Li Y-K (2007) Mutagenesis and mechanistic study of a glycoside hydrolase family 54 α-L-arabinofuranosidase from *Trichoderma koningii*. Biochem J 401:551–558

Wang Y, Wanga Z, Xu Q, Dua G, Hua Z, Liu L, Li J, Chen J (2009) Lowering induction temperature for enhanced production of polygalacturonate lyase in recombinant *Pichia pastoris*. Process Biochem 44:949–954

Yanay T, Sato M (2000) Purification and characterization of a novel α-L-arabinofuranosidase from *Pichia capsulata* X91. Biosci Biotechnol Biochem 64(6):1181–1188