



**UNIVERSITÀ DEGLI STUDI DELLA TUSCIA
VITERBO
Dipartimento di Scienze Ambientali**

**DOTTORATO DI RICERCA IN
“GENETICA E BIOLOGIA CELLULARE” - XVIII CICLO**

**Proteomica come piattaforma di indagine dalle molteplici
potenzialità applicative**

Docente Guida: Prof. Lello Zolla

Coordinatore: Prof. Luigi Bosco

Dottorando: Corrado Ciambella

"Siamo nani arrampicati sulle spalle di giganti. Così vediamo di più e più lontano di loro, non perché la nostra vista sia più acuta o la statura più alta, ma perché ci solleviamo in aria con tutta la loro altezza gigantesca..."

Berardo di Chartres

*Alle mie nonne,
Annunziata, Giacinta ed Alda*

Indice

INDICE		i
SCOPO DELLA TESI		1
PARTE I	“INTRODUZIONE”	4
CAPITOLO 1	LA PROTEOMICA	5
1.1	<i>Cos'è la proteomica?</i>	5
1.2	<i>Il binomio Proteomica/Spettrometria di massa</i>	5
1.2.1	<i>Protemica d'espressione</i>	5
1.2.2	<i>Protemica differenziale e quantificazione</i>	9
1.2.3	<i>Modificoma: proteomica e studio delle modificazioni post-traduzionali (PTMsoica)</i>	17
CAPITOLO 2	SPETTROMETRIA DI MASSA: TECNICA D'ELEZIONE PER L'INDAGINE PROTEOMICA	20
2.1	<i>Formazione degli ioni in fase gassosa</i>	21
2.1.1	<i>Ionizzazione Laser Assistita da Matrice (MALDI)</i>	22
2.1.2	<i>Ionizzazione per Elettrospray (ESI)</i>	26
2.2	<i>Analizzatore</i>	28
2.2.1	<i>Analizzatore a tempo di volo (TOF)</i>	28
2.2.1.1	<i>Principi dell'analizzatore TOF</i>	29
2.2.1.2	<i>Evoluzione dell'analizzatore a tempo di volo</i>	30
2.2.2	<i>Analizzatore a quadrupolo (Q)</i>	33
2.2.3	<i>Spettrometria di massa tandem (MS/MS)</i>	35
2.2.3.1	<i>Dissociazione indotta/attivata da collisione (CID/CAD)</i>	35

2.2.3.2	<i>Nomenclatura per la frammentazione peptidica</i>	36
2.2.4	<i>Spettrometri di massa ibridi quadrupolo TOF (Q-TOF)</i>	38
2.2.5	<i>Analizzatori di massa ad intrappolamento ionico</i>	39

CAPITOLO 3	IDENTIFICAZIONE DELLE PROTEINE	43
------------	--------------------------------	----

3.1	<i>Identificazione “top-down”</i>	43
3.2	<i>Identificazione “bottom-up”</i>	44
3.2.1	<i>Peptide Mass Fingerprinting (PMF)</i>	44
3.2.2	<i>Digestione in-gel</i>	45
3.2.3	<i>Software di interrogazione delle banche dati</i>	45
3.3	<i>Sequenziamento peptidico</i>	46
3.3.1	<i>Sequenziamento in silico (identificazione probabilità mediata)</i>	46
3.3.2	<i>Sequenziamento de novo</i>	47
3.3.3	<i>Frammentazione chimicamente assistita (CAF)</i>	48
3.3.4	<i>Derivatizzazione con 4-sulfofenil-isotiocianato (SPITC)</i>	50

CAPITOLO 4	BIBLIOGRAFIA	53
------------	--------------	----

PARTE II	“UTILIZZO DI UNA STRATEGIA DI INDAGINE DI TIPO PROTEOMICO PER LO STUDIO DELL’APPARATO FOTOSINTETICO NELLE PIANTE SUPERIORI”	68
-----------------	--	-----------

<i>Proteomics</i> (2005) 5: 746-757	69
-------------------------------------	----

PARTE III	“ELETTROFORESI BIDIMENSIONALE E SPETTROMETRIA DI MASSA NELL’IDENTIFICAZIONE DI PROTEINE CARATTERIZZANTI LE SPECIE <i>Coffea arabica</i> (var. Colombia) E <i>Coffea canephora</i> (var. Indiano Robusta)”	81
------------------	--	-----------

<i>Proteomics</i> (2005) 5: 710-718	82
-------------------------------------	----

**“ANALISI DELLE PROTEINE S-NITROSILATE ESPRESSE
IN *Arabidopsis thaliana* NELLA RISPOSTA
ALL’INFEZIONE DEL PATOGENO *Pseudomonas syringae***

PARTE IV	pv. <i>tomato</i>”	91
	ABSTRACT	92
CAPITOLO 1	INTRODUZIONE	93
1.1	<i>L’ossido nitrico (NO)</i>	93
1.2	<i>La S-nitrosilazione</i>	95
1.3	<i>Metodologie di rivelazione per le proteine S-nitrosilate</i>	98
1.4	<i>Produzione di NO nelle piante</i>	100
1.5	<i>Ruolo dell’ossido nitrico nei sistemi vegetali</i>	102
1.6	<i>NO nella risposta di ipersensibilità ai patogeni</i>	103
CAPITOLO 2	MATERIALI E METODI	106
2.1	<i>Materiale vegetale</i>	106
2.2	<i>Materiale microbiologico</i>	106
2.3	<i>Infiltrazione di foglie di <i>A. thaliana</i> con <i>Pst AvrB</i></i>	106
2.4	<i>Estrazione proteica</i>	106
2.5	<i>Purificazione per immunoprecipitazione delle proteine S-nitrosilate</i>	107
2.6	<i>Elettroforesi bidimensionale (2-DE)</i>	108
2.7	<i>Acquisizione ed analisi delle immagini</i>	108
2.8	<i>Digestione triptica delle proteine in-gel</i>	110
2.9	<i>Spettrometria di massa MALDI-TOF/TOF</i>	110
CAPITOLO 3	RISULTATI	112
	DISCUSSIONE	122

CAPITOLO 4		
CAPITOLO 5	BIBLIOGRAFIA	126
PARTE V	“SEQUENZIAMENTO DE NOVO DELLE SUBUNITÀ α E β DELLA FICOCIANINA DI <i>Aphanizomenon flos-aquae</i> MEDIANTE SPETTROMETRIA DI MASSA TANDEM MALDI- TOF/TOF”	136
	ABSTRACT	137
CAPITOLO 1	INTRODUZIONE	138
1.1	<i>Aphanizomenon flos-acquae</i>	138
1.2	<i>Sistema antenna nei cianobatteri: i Ficobilisomi</i>	140
CAPITOLO 2	MATERIALI E METODI	146
2.1	<i>Purificazione delle PC</i>	146
2.2	<i>SDS-PAGE</i>	146
2.3	<i>Digestione triptica delle proteine in-gel</i>	146
2.4	<i>Determinazione della massa intatta della proteina: MALDI-TOF-MS</i>	147
2.5	<i>Spettrometria di massa MALDI-TOF/TOF: Peptide Mass Fingerprinting (PMF) e De Novo sequencing</i>	148
2.6	<i>Derivatizzazione chimica con 4-sulfofenil-isotiocianato (SPITC)</i>	149
CAPITOLO 3	RISULTATI	150
CAPITOLO 4	DISCUSSIONE	162

CAPITOLO 5	BIBLIOGRAFIA	165
	RINGRAZIAMENTI	169
	PUBBLICAZIONI	170
	COMUNICAZIONI A CONGRESSI	170

SCOPO DELLA TESI

La proteomica, rappresenta lo studio sistematico delle molteplici e differenti proprietà delle proteine, utilizzando una varietà di approcci che sinergicamente hanno lo scopo di permettere una descrizione dettagliata della struttura, della funzione e dei meccanismi di controllo dei sistemi biologici. Il progredire delle metodiche scientifiche e delle tecnologie nelle ultime decadi, ha catalizzato un'estensione degli scopi connessi agli studi in campo biologico spostando l'attenzione dalle riduttive analisi biochimiche a carico delle singole proteine ad un esame attento ed approfondito esteso all'intero proteoma. La proteomica, nasce come naturale evoluzione della chimica delle proteine che rappresentava uno dei capisaldi nello studio della biologia durante gli anni ottanta. Lo sviluppo esponenziale delle scienze tecnologiche in quegli anni, ha portato i ricercatori a disporre di un numero sempre più crescente di frecce nella loro faretra: tecniche di clonaggio dei geni e di sequenziamento proteico, metodiche di analisi per l'espressione genica, ottimizzazione delle tecniche separative elettroforetiche, evoluzione della spettrometria di massa, ampliamento dei database di sequenza e dei relativi programmi di interrogazione degli stessi. L'insieme di tutte queste componenti ha fornito ai chimici delle proteine il cavallo di Troia per uno studio su larga scala delle molecole proteiche. E' in questo contesto storico-scientifico che nel 1994, al "I 2DE meeting" a Siena, viene coniato il termine "Proteoma". Il proteoma viene definito come il complemento proteico al genoma e di conseguenza lo studio dello stesso viene indicato con il termine "Proteomica" [1].

La proteomica ed altre metodologie di indagine complementari (genomica, microarray, metabolomica), costituiscono i componenti essenziali di una nuova ed emergente tipologia di studio che viene oggi indicata con il termine "systems biology" [2]. In quest'ottica, la "systems biology" si interessa di studiare un determinato sistema biologico attraverso delle analisi sistematiche e quantitative di tutte le componenti che concorrono alla definizione dello stesso sistema preso in esame. Questo approccio, mira alla descrizione estensiva di un sistema biologico attraverso l'integrazione di dati di tipo differente, con lo scopo, in futuro, di poter giungere a realizzare delle simulazioni computazionali che mimino il comportamento dei sistemi biologici più complessi. Alla base di questo nuovo modo di intendere lo studio della biologia è sottesa la considerazione che per una data specie, il numero delle possibili biomolecole e la loro organizzazione in vie

metaboliche e processi biologici è sì ampio, ma finito. Da ciò, teoricamente, si evince che i sistemi biologici operanti in una determinata specie possono essere descritti nella loro interezza se si ottiene una sufficiente densità di osservazioni di tutti gli elementi che lo costituiscono.

E' in quest' ottica che si può capire l'importanza dell'evolversi della proteomica stessa e la sua specializzazione nel cercare di rispondere a domande di tipo differente e nel descrivere accuratamente diversi aspetti del complesso mondo delle molecole proteiche. Maggiore e più completa è la caratterizzazione di una proteina in termini di sequenza, struttura e funzione, maggiore è il contributo che si riesce a dare alla descrizione d'insieme del sistema biologico ove tale prodotto genico si trova ad assolvere il proprio ruolo.

Questo lavoro di tesi nasce quindi come viaggio personale del candidato alla scoperta del mondo della proteomica, delle sue potenzialità e delle possibili applicazioni. Il primo step sarà costituito da un'indagine proteomica di tipo "classico" volta all'identificazione delle proteine di membrana presenti nei tilacoidi di orzo (*Hordeum vulgare*). In questo tipo di ricerca, si utilizzerà un approccio alternativo per lo studio dei complessi fotosintetici che verranno risolti attraverso un'elettroforesi bidimensionale (2-DE) in cui si sostituirà all'isoelettrofocusing (IEF) in prima dimensione un Blue Native gel (BN) non denaturante. Questa tecnica, permette di ottenere informazioni sull'aggregazione delle proteine coinvolte nel processo fotosintetico in specifici complessi e/o super-, sub-complessi. Una seconda dimensione denaturante utilizzando gel di poliacrilammide in presenza di sodio dodecil-solfato (SDS-PAGE), ci consentirà di risolvere i complessi nelle singole componenti che verranno poi identificate con l'ausilio della spettrometria di massa (MS).

Successivamente in questo percorso di studio, si valuterà la possibilità di utilizzare l'approccio proteomico "classico" con finalità differenti rispetto alla semplice identificazione proteica. In questa direzione, l'utilizzo dell'elettroforesi bidimensionale IEF/SDS-PAGE e della spettrometria di massa verranno adottate per realizzare uno studio filogenetico preliminare, attraverso l'ausilio di tecniche quantitative di analisi proteomica differenziale; l'intento è quello di individuare quali siano le catene polipeptidiche (caratterizzanti) espresse in un'unica varietà di caffè (*Coffea arabica*/*Coffea Robusta*) o quali proteine risultino sotto- e/o sovra-espresse quando si esegue tra le diverse varietà un'analisi di tipo comparativo.

Nella continua specializzazione che si viene a delineare nell'ambito della proteomica, una delle sottobranche d'indagine che oggi assume sempre un maggiore interesse è quella che a volte viene anche indicata con il termine di PTMs-omica, dall'inglese "Post Traslational Modifications". Come indicato dal nome stesso, tale branca si occupa dello studio del "Modificoma", ovvero delle modifiche post-traduzionali a carico delle molecole proteiche: identificazione delle molecole modificate, descrizione del tipo di modifica ed eventuale definizione del sito modificato. Lo studio delle modificazioni post-traduzionali rappresenta a tutt'oggi uno degli aspetti più avvincenti della proteomica considerando che si conoscono più di duecento modificazioni a cui le proteine possono essere sottoposte (acilazioni, fosforilazioni, glicosilazioni, nitrosilazioni, processamenti proteolitici, ecc.) e che in ultima analisi queste giocano un ruolo fondamentale nella definizione della funzionalità del prodotto proteico. A tal riguardo, in questo lavoro di tesi ci si occuperà dello studio della S-nitrosilazione a carico delle proteine di *Arabidopsis thaliana* indotta a seguito dell'interazione con un ceppo avirulento del patogeno batterico *Pseudomonas syringae* pv. tomato. Le proteine modificate saranno dapprima selettivamente isolate utilizzando il cosiddetto metodo "biotin-switch" poi in seguito separate attraverso 2-DE ed infine caratterizzate con l'ausilio delle tecniche di spettrometria di massa.

L'ultimo obiettivo di questo lavoro è rappresentato dalla caratterizzazione di due ficobiliproteine, le ficocianine (PC) α e β del cianobatterio *Aphanizomemon flos-aquae*, coinvolte nel processo di assorbimento della luce da parte dell'apparato fotosintetico. A tal riguardo, le proteine preventivamente separate grazie all'ausilio di un gradiente di saccarosio e SDS-PAGE, verranno utilizzate per le analisi in spettrometria di massa. L'indagine prevederà la determinazione del peso molecolare intatto dei prodotti genici ed il sequenziamento "de-novo" degli stessi attraverso l'utilizzo della tecnica di derivatizzazione chimica che fa uso dello SPITC (4-sulfofenil-isotiocianato) allo scopo di semplificare gli spettri e la loro conseguente interpretazione.

PARTE I

“INTRODUZIONE”

CAPITOLO 1

LA PROTEOMICA

1.1 Cos'è la proteomica?

“Proteomics is not just a mass spectrum of a spot on a gel” [3].

La proteomica nasce come scienza che mira ad indagare e stabilire l'identità, la quantità, la struttura e le funzioni biochimiche e cellulari di tutte le proteine in un organismo, organo od organello, descrivendo come queste proprietà varino nello spazio, nel tempo o in un determinato stato fisiologico (Fig 1.1). La ricerca nelle banche dati, che permettono l'accesso alla letteratura scientifica, mostra chiaramente come il termine proteomica possa significare cose molto diverse per persone differenti. Nella sua espressione più classica, il termine proteomica indica lo studio condotto attraverso la separazione di un determinato campione proteico per mezzo di elettroforesi bidimensionale (2-DE), seguita dall'identificazione dei singoli spot mediante l'ausilio della spettrometria di massa (MS) [4-6]. Recentemente lo sviluppo di una proteomica di tipo funzionale e strutturale ha portato ad un ampliamento di intenti conducendo la ricerca un passo più vicino ad una migliore comprensione dei processi e delle funzioni cellulari. [7].

1.2 Il binomio Proteomica/Spettrometria di massa

Lo studio della proteomica, attraverso l'utilizzo delle tecniche di spettrometria di massa, ha permesso nel corso del suo sviluppo di poter affrontare in maniera ottimale tre rilevanti tipi di problematiche biologiche:

- a) identificazione delle proteine
- b) quantificazione ed analisi dei profili di espressione proteici
- c) studio delle modifiche post-traduzionali (PTMs) a carico dei prodotti genici [8]

1.2.1 Proteomica d'espressione

La proteomica d'espressione rappresenta l'analisi su larga scala dell'espressione genica condotta a livello proteico. La nascita di questo approccio di studio è il risultato dello sviluppo crescente delle tecnologie analitiche volte all'identificazione delle molecole proteiche preventivamente separate, del successo dei programmi di sequenziamento e del boom evolutivo che ha investito il settore

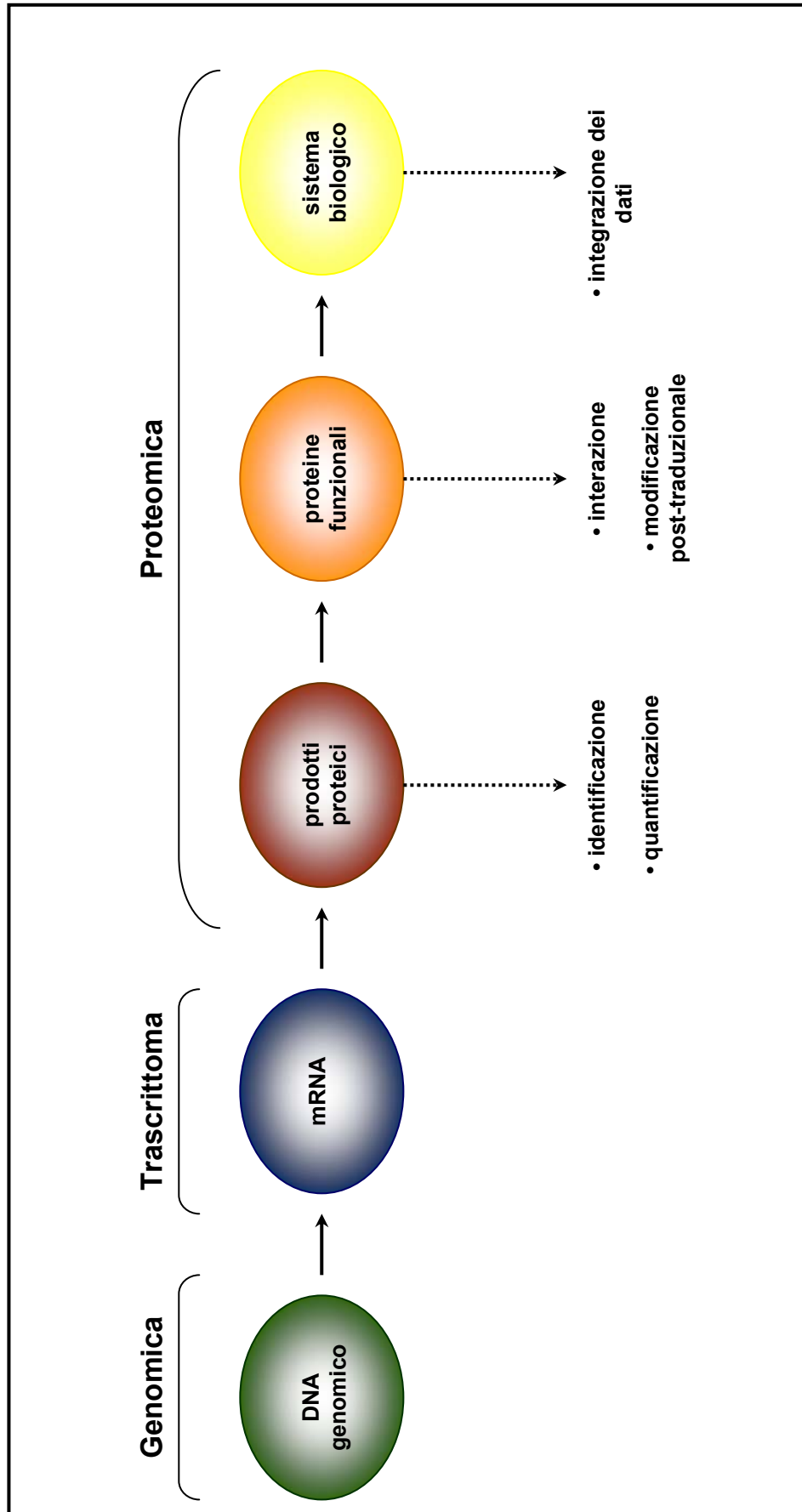


Figura 1.1 Dalla genomica alla proteomica. Integrazione delle informazioni dal genoma al proteoma per una comprensione migliore dei sistemi biologici. Figura adattata da [3].

della bioinformatica. Tradizionalmente, la proteomica d'espressione è caratterizzata da una procedura consecutiva che prevede due step principali: la separazione delle catene polipeptidiche mediante 2-DE e l'analisi delle stesse via spettrometria di massa. L'elettroforesi bidimensionale permette di separare le proteine utilizzando due parametri ortogonali. In prima dimensione, la separazione avviene in base alla carica della molecola attraverso un processo di isoelettrofocalizzazione (Isoelettrofocusing, IEF), per cui i prodotti genici sono distinti grazie alle differenze dei punti isoelettrici (pIs) che li caratterizzano. La seconda dimensione si realizza grazie all'ausilio di un gel di poliacrilamide in presenza di sodio dodecil-solfato (SDS-PAGE) e garantisce che le proteine siano invece risolte in base alla differenza di peso molecolare. La 2-DE possiede molteplici caratteristiche e proprietà che hanno reso difficile la ricerca di una valida alternativa che possa sostituirla negli studi volti alla mappatura del proteoma. A tutt'oggi, non esiste un'altra tecnica che possa fornire informazioni ricche e dettagliate su una particolare proteina in esame in termini di quantità relativa, PTMs, pI, peso molecolare e solubilità. L'applicazione di questa tecnologia, ha permesso a diversi gruppi di poter iniziare un'analisi globale delle proteine così da permettere la costruzione sistematica di banche dati contenenti mappe bidimensionali. Questi database consentono di visualizzare i dati relativi alle catene polipeptidiche identificate semplicemente cliccando sullo spot di interesse e permettono di eseguire delle ricerche utilizzando il nome della proteina, parole chiavi, masse molecolari, punti isoelettrici, organelli o compartimenti cellulari.

Nonostante l'efficienza di separazione, fino a 10.000 spot proteici utilizzando gradienti immobilizzati di pH in un range ristretto (IPGs) [9], la 2-DE ha le sue limitazioni. Queste restrizioni sono legate all'elevata diversità chimica delle proteine e al loro differente livello di espressione nella cellula [10]. Le più recenti tecniche di colorazione fluorescente, permettono di coprire un range dinamico di quattro o cinque ordini di grandezza mentre il proteoma appare spingersi anche oltre i dodici ordini di grandezza quando si mettono a confronto la più abbondante proteina presente nel plasma, l'albumina, ed una singola copia di una molecola segnale [10]. Da ciò, ne consegue che la proteomica che si avvale della 2-DE è in qualche modo prevenuta contro l'identificazione delle proteine meno abbondanti [11-13]. Un'altra difficoltà è costituita dalla capacità di ottenere un corretto rilevamento di proteine idrofobiche, basiche, di dimensioni piccole o molto grandi. Molecole proteiche scarsamente solubili (come le proteine di membrana) sfuggono

alla rivelazione 2-DE mediata a causa proprio della loro insolubilizzazione durante il processo di preparazione del campione, durante l'ingresso nel gel di isoelettrofocusing o ancora durante il processo stesso di isoelettrofocalizzazione, quando le molecole raggiungono il proprio punto isoelettrico. Inoltre, le piccole proteine (minori di 8 kDa) vanno perse durante l'IEF mentre quelle troppo grandi (al di sopra dei 200 kDa) fanno il loro ingresso nel gel con un'efficienza estremamente bassa. Infine occorre considerare come le catene polipeptidiche con un elevato grado di acidità o basicità possono causare problemi, poichè il loro valore di pI eccede rispetto ai range di pH comunemente disponibili per la realizzazione della prima dimensione.

La consapevolezza dell'esistenza di tali limitazioni ha spinto la ricerca concernente le tecnologie applicate alla proteomica a cercare di sviluppare delle metodiche alternative che non siano basate sull'applicazione dell'elettroforesi bidimensionale. I primi risultati promettenti si ebbero quando Hunt e collaboratori [14] dimostrarono la capacità della cromatografia liquida combinata con la spettrometria di massa (LC-MS/MS) di poter analizzare complesse miscele peptidiche attraverso un approccio che non prevedeva l'utilizzo di gel. Poco dopo, Yates ed i suoi colleghi [15] svilupparono una sistema di separazione multidimensionale in fase liquida interfacciato con uno spettrometro di massa definendolo con il nome di "Multidimensional Protein Identification Technology", o MudPIT. Tale metodica, alla quale ci si riferisce anche con il termine di "shotgun proteomics" [16], mostra la capacità di rendere possibile la catalogazione di centinaia di migliaia di componenti generate da un singolo esperimento. Tutto questo è esemplificato dall'identificazione di 1.484 proteine di lievito, realizzata utilizzando una strategia automatizzata in cui le proteine sono separate attraverso un sistema 2-D LC capillare (scambio ionico/fase inversa) accoppiato on line con uno sistema che permette di realizzare analisi di MS² [17]. Le performance di questo tipo di approccio sono di lunga superiori a quelle che si possono ottenere utilizzando invece un tipo di approccio 2-DE mediato; dalla letteratura si evince infatti che nel migliore dei casi si sono identificate 502 proteine appartenenti al batterio *Haemophilus influenzae* [18].

Un ulteriore approccio per lo studio del proteoma è stato sviluppato da Smith e collaboratori [19-22] i quali analizzarono un digerito proteolitico *in toto*, utilizzando un sistema di cromatografia capillare ad altissima risoluzione (ultra-high), in combinazione con uno spettrometro di massa a risonanza ionica

ciclotronica in trasformata di Fourier (FT-ICR-MS). L'uso delle apparecchiature FT-ICR permette di raggiungere un'accuratezza di massa di poche femtomoli. Tutto questo consente di identificare le proteine basandosi sulla massa di un singolo frammento proteolitico, evitando il bisogno di esperimenti di massa tandem che determinano un notevole aumento dei tempi di analisi. Un ben determinato frammento proteolitico può, in quest'ottica, essere considerato come un biomarker o come alternativamente si è soliti dire, un "mass-tag" che indica la presenza in un campione della sua proteina di origine. L'accoppiamento on-line di cromatografia liquida ad alta risoluzione con la spettrometria FT-ICR, potenzialmente, determina le desiderate alte prestazioni, sensibilità e range dinamico per un'identificazione su larga scala delle proteine. D'altro canto, occorre però considerare che seppure la "shotgun proteomics" ha notevoli vantaggi, presenta però il limite di non consentire il più delle volte la rivelazione e l'analisi di tutti i peptidi presenti nel campione preso in esame.

1.2.2 Proteomica differenziale e quantificazione

L'analisi quantitativa e il confronto dell'espressione proteica tra un sistema biologico in condizioni normali ed uno in condizioni patologiche, o il monitoraggio delle stesse caratteristiche durante le diverse fasi di sviluppo del sistema stesso, rappresenta uno degli obiettivi principali perseguiti dall'indagine proteomica [22]. La metodica correntemente più usata per monitorare i cambiamenti di espressione proteica è la 2-DE, attraverso la quale le catene polipeptidiche sono visualizzate e quantificate per mezzo delle procedure di colorazione. Il blue Coomassie è stato di gran lunga il colorante maggiormente utilizzato a causa della semplicità d'uso e per la sua compatibilità con altre tecniche analitiche come il sequenziamento di Edman e la spettrometria di massa. Sfortunatamente, la sua relativa sensibilità (approssimativamente 1 µg) limita la sua applicazione per lo studio del proteoma. I coloranti commerciali a disposizione, con il più elevato grado di sensibilità, sono il silver-staining ed i coloranti fluorescenti i quali permettono di rilevare nanogrammi di proteine presenti nel gel. Ad ogni modo, il silver-staining presenta un ristretto range dinamico e proteine differenti mostrano diverse caratteristiche di colorazione compromettendo in questo modo l'accuratezza durante la quantificazione delle molecole proteiche. Diversa è la situazione quando si esaminano i coloranti fluorescenti come il Sypro Ruby ed il Sypro Orange che offrono invece ulteriori vantaggi. Il Sypro Ruby ha un ampio range dinamico (oltre i tre ordini di

grandezza), permettendo una migliore quantificazione relativa delle catene polipeptidiche separate quando lo si confronta con il silver-staining. In aggiunta, si osserva una scarsa variabilità nella capacità di colorare le diverse proteine, dovuta al fatto che il Sypro interagisce con le molecole di SDS e non con la catena proteica.

Un approccio per una quantificazione differenziale è invece quello che prevede l'uso della "Difference Gel Electrophoresis" (DIGE) che utilizza la formazione di legami covalenti tra gli ϵ ammino-gruppi della lisina con differenti coloranti fluorescenti. Quando si esegue questa procedura, le proteine di due o tre popolazioni di campioni sono marcate differentemente con dei coloranti fluorescenti (succinimidil esteri o diversi coloranti derivati dalla cianina Cy2, Cy3, Cy5). Questo, rende possibile utilizzare un solo gel per la separazione proteica e per la quantificazione delle differenze tra i campioni. Il grande vantaggio di tale tecnica è quello di minimizzare le variazioni degli spot da gel a gel [23, 24].

Poichè la quantificazione attraverso la 2-DE si basa sull'assunzione che in ciascuno spot sia presente una proteina, i fenomeni di co-migrazione possono compromettere quest'analisi. Inoltre, occorre ricordare che una rappresentazione completa del proteoma è comunque impossibile. Quindi, seppur considerando l'elevato grado di "maturità" della tecnica e le ineguagliate performance per la separazione di complessi pattern di proteine differenzialmente modificate e processate, e seppur la 2-DE sia in continua evoluzione, si è fatta sentire la necessità da parte della comunità scientifica internazionale di cercare delle metodiche alternative.

Un passo significativo nel campo della proteomica differenziale è stato fatto quando si è sviluppato il procedimento di marcatura isotopica seguito da analisi quantitativa attraverso spettrometria di massa. Questo approccio si basa sulla marcatura differenziale del campione controllo e di quello sperimentale, il loro mescolamento e la misura delle differenze relative dell'espressione proteica tra i due MS-mediata. A tutt'oggi, esistono molteplici tecniche di marcatura e ciascuna può essere utilizzata in maniera peculiare per marcare le proteine durante un diverso step di preparazione del campione (Fig. 1.2). Brevemente diamo un'occhiata alle metodiche principali che hanno destato maggiore interesse nello sviluppo delle tecniche di quantizzazione.

La marcatura *in vivo*, attraverso la crescita delle cellule in un terreno arricchito in ^{15}N è stata introdotta da Chait e collaboratori [25]. Le proteine

provenienti da un campione controllo e da uno marcato sono state combinate, separate attraverso SDS-PAGE ed infine quantificate utilizzando la spettrometria di massa. Questo metodo è stato in seguito modificato attraverso la marcatura di specifici residui amminoacidici invece che marcare l'intero pool di amminoacidi con

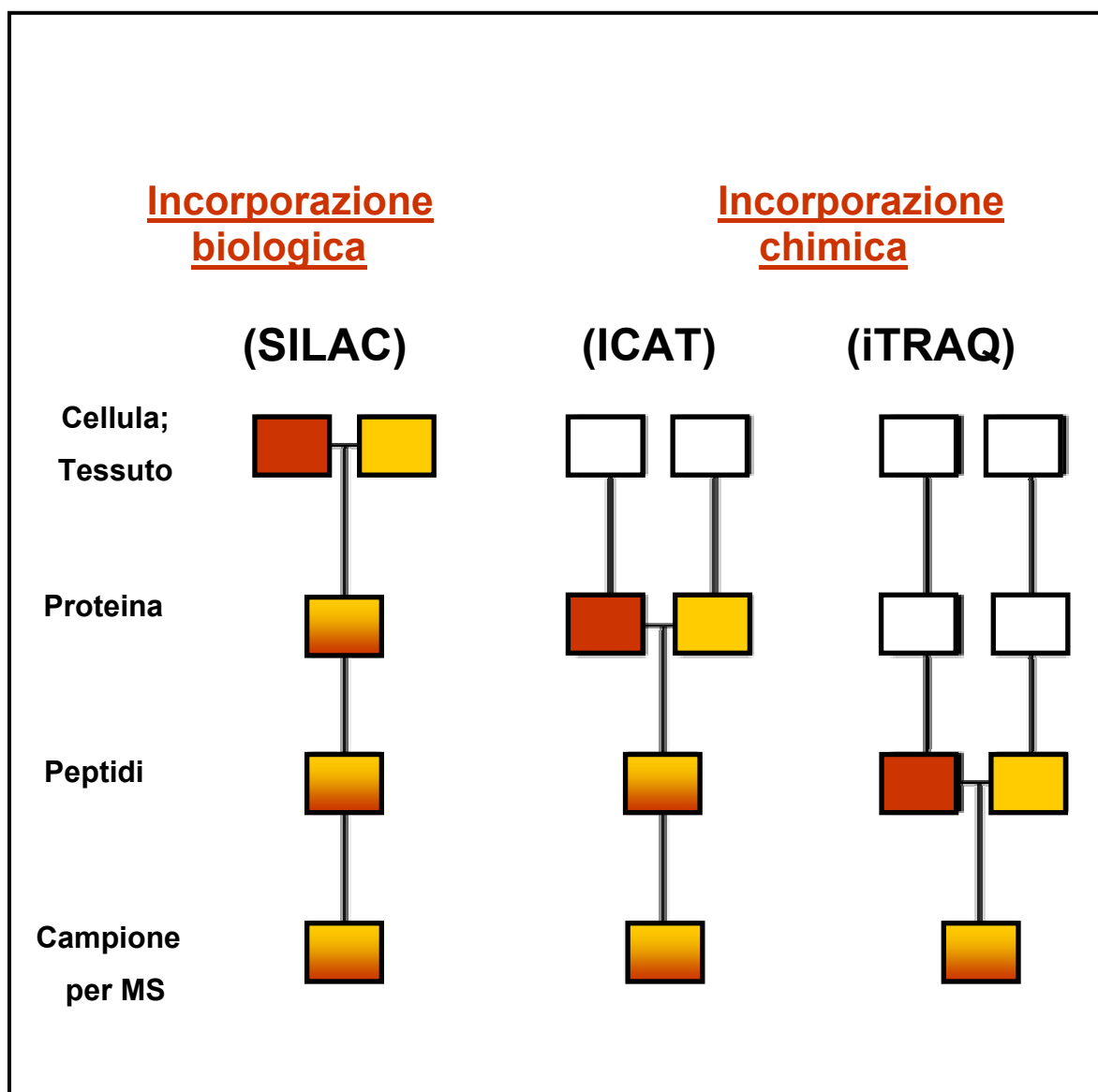


Figura 1.2 Illustrazione mostrante alcune delle molteplici tecniche di marcatura del campione per la quantificazione MS mediata. La marcatura può essere introdotta attraverso procedimenti biologici o chimici e durante diversi step di preparazione del campione.

¹⁵N. Tale modifica è stata descritta nella definizione di un nome *ad hoc* per tale metodica indicata come "Stable Isotope Labelling by Amino acid in Cell culture" (SILAC) [26]. In un primo tempo la procedura è stata introdotta per un'accurata e

specifica identificazione proteica [26] e solo in un secondo step la si è utilizzata per la proteomica quantitativa [27]. Diversi amminoacidi sono stati utilizzati come, $^2\text{H}_3$ -leucina e ^{15}N -metionina. Il metodo di marcatura *in vivo* è un mezzo potente e ben consolidato, grazie al fatto che le due popolazioni di cellule differentemente marcate vengono combinate prima dei vari step di processamento del campione (Fig. 1.3), permettendo quindi di eliminare tutte le variabili sperimentali associate alla lisi della cellula, alla separazione e all'analisi MS. I principali limiti di questa metodica risiedono nell'elevato costo del materiale utilizzato per la marcatura, nei lunghi tempi di processamento, nella possibile non prevedibilità legata all'incorporazione della marcatura e nella sua non applicabilità ai campioni tissutali. Ciò ha determinato la ricerca di protocolli alternativi per un'incorporazione chimica della marcatura piuttosto che biologica.

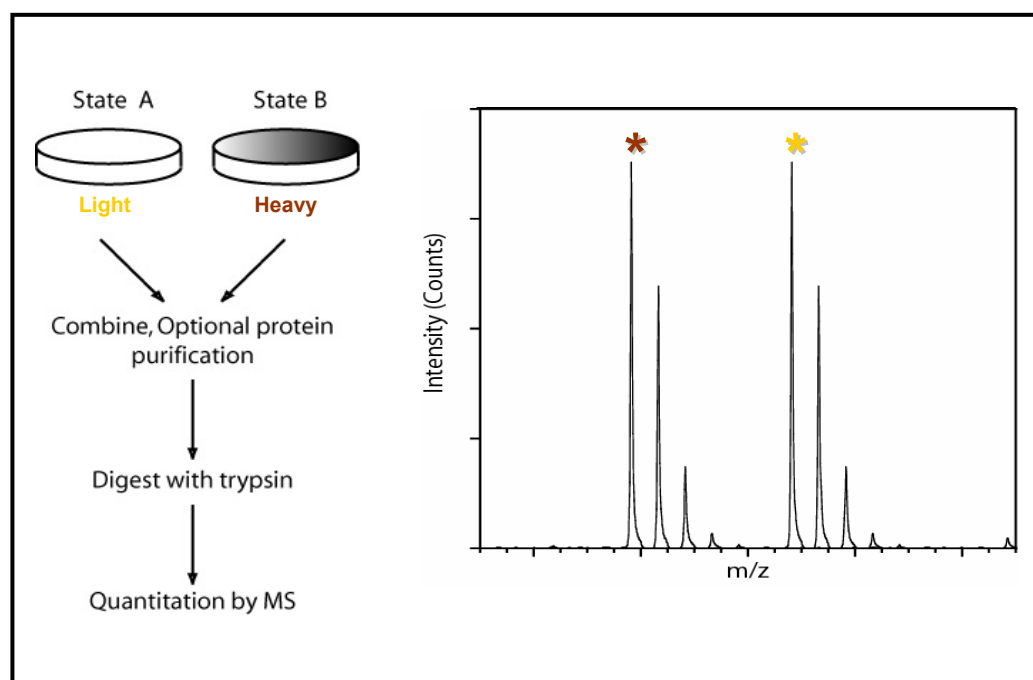


Figura 1.3 Marcatura dei campioni mediante SILAC: due popolazioni di cellule sono marcate attraverso incorporazione biologica utilizzando amminoacidi contenenti isotopi differenti. Ogni peptide contenente la differente marcatura dà origine ad un doppietto nello spettro di massa. La stima dei rapporti tra le intensità dei segnali del doppietto fornisce una quantificazione relativa della proteina, dal quale il peptide deriva, nella popolazione originale.

I campioni proteici, possono infatti anche essere marcati dopo la loro purificazione dal materiale di partenza ed essere poi messi a confronto. In tal caso il metodo più associato è quello indicato come “Isotope-Coded Affinity Tagging”

(ICAT) [28]. L'ICAT prevede la derivatizzazione della catena laterale di un residuo cisteinico con un reagente isotopicamente pesante (deuterato) per un campione e quella isotopicamente leggera (reagente non deuterato) per il secondo campione. Dopo la marcatura i campioni vengono combinati e digeriti per mezzo di una endoproteasi specifica. La peculiare struttura del reagente di derivatizzazione (Fig. 1.4), che possiede un tag per la biotina, garantisce che dopo la proteolisi i peptidi contenenti la marcatura possano essere selettivamente isolati attraverso una cromatografia di affinità basata sull'interazione avidina/biotina, permettendo una notevole riduzione della complessità del sistema.

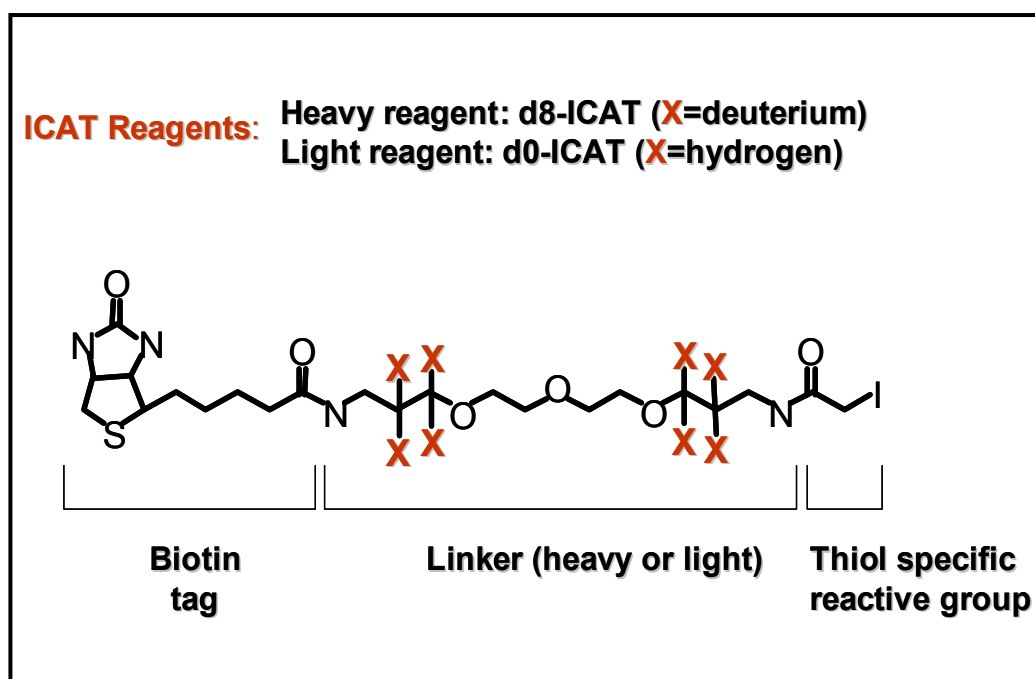


Figura 1.4 Struttura chimica del reagente ICAT. Il gruppo iodo-acetile reagisce selettivamente con i residui cisteinici (Cys) delle catene proteiche.

Le coppie peptidiche marcate con il reagente ICAT rispettivamente leggero e pesante, sono chimicamente identiche e servono quindi come standard interni per un'accurata quantificazione. In spettrometria di massa il rapporto tra le intensità dei picchi del componente a massa più bassa e di quello a massa più alta di queste coppie peptidiche fornisce l'abbondanza relativa dei peptidi e conseguentemente l'abbondanza delle proteine nelle cellule di origine. Il principale svantaggio di questa strategia di marcatura chimica, risiede nel fatto che solo le proteine contenenti cisteina vengono derivatizzate e marcate. In aggiunta la metodica

riduce il numero delle identificazioni delle proteine che vengono post-traduzionalmente modificate.

La terza strategia di marcatura che è stata investigata è quella che avviene a carico dei peptidi durante o dopo la digestione proteolitica. L' ^{18}O , ad esempio, può essere incorporato nei peptidi a livello dell'acido carbossilico all'estremità C-terminale in una reazione catalizzata enzimaticamente durante il processo di digestione. [29]. Seguendo questa procedura, un set di proteine viene digerito in H_2^{18}O , mentre un altro in presenza di H_2^{16}O , mischiando poi i campioni dopo la reazione. Seppure il processo di incorporazione avviene abbastanza semplicemente, la quantificazione proteolitica non è accurata per il fatto che la marcatura avviene nell'ultimo step che precede l'analisi MS. Inoltre occorre considerare l'elevato costo dell'agente utilizzato per la marcatura stessa.

Non di meno i peptidi posso anche essere marcati dopo la digestione proteolitica della catena polipeptidica utilizzando protocolli diversi che permettono alternativamente di derivatizzare le porzioni N-terminali, C-terminali o le catene laterali di diversi amminoacidi.

I metodi di marcatura chimica appena descritti, ICAT e marcatura con ^{18}O , consentono una quantificazione del campione MS-basata. Esiste però una nuova strategia che fa uso di un reagente commerciale (Applied Biosystems) indicato come iTRAQ che permette una quantificazione MS/MS-basata. L'iTRAQ (Fig. 1.5) è costituito da tre porzioni ben distinte:

- a) peptide reactive group: tale porzione ammino-specifica permette la derivatizzazione dei peptidi all'N-terminale
- b) reporter: è la porzione carica della molecola che mantiene lo stato di carica e da origine ad uno ione di marcatura specifico nello spettro MS/MS in una regione "pulita" dello stesso. Il reporter può avere quattro valori di massa differenti (dai 114 ai 117 Da) consentendo la marcatura differenziale di quattro campioni diversi e non solo di due come per l'ICAT. Inoltre assolve un ruolo fondamentale nel garantire un'efficiente ionizzazione dell'analita
- c) balance: ha il compito, come indica il nome, di bilanciare le differenze di massa dei vari reporter consentendo al binomio reporter/balance di costituire un tag isobarico di 145 Da per tutti i differenti campioni. Inoltre dà origine ad una perdita neutra negli esperimenti di MS/MS.

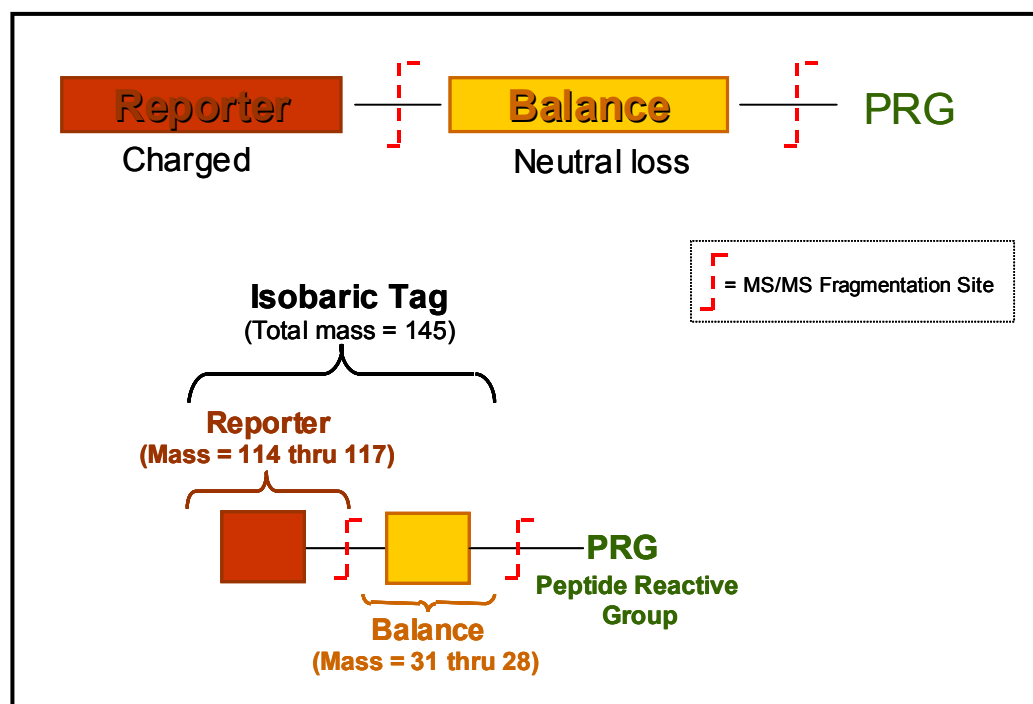


Figura 1.3 Struttura schematizzata del reagente iTRAQ.

L'iTRAQ viene utilizzato per marcare i campioni, per i quali si vuole valutare la differenza nell'espressione proteica, dopo aver effettuato la digestione proteolitica. Diversi campioni vengono marcati con molecole del reagente contenente reporter di peso differente. Avvenuta la reazione di derivatizzazione peptidica i campioni vengono miscelati insieme ed analizzati in MS. Peptidi provenienti dalla stessa proteina, espressa differenzialmente nei vari campioni analizzati e legati a reagenti iTRAQ con diversi reporter, risultano in uguali valori di massa negli spettri MS grazie al bilanciamento dovuto alla presenza dei differenti balance. E' già a questo livello che si evidenzia un primo vantaggio nell'utilizzo di tale tipo di tecnica legato ad un incremento del segnale per un peptide esaminato, dovuto alla combinazione dei segnali provenienti da ciascun differente campione.

La frammentazione di un peptide di interesse derivatizzato con l'iTRAQ, determina la rottura dei legami che mantengono insieme le tre porzioni del reagente, liberando in forma ionica i diversi reporter che vengono rivelati nello spettro MS/MS a valori di m/z compresi tra 114 e 117 (Fig. 1.6). I valori dei reporter sono stati scelti in maniera che questi ioni andassero a cadere in una zona dello spettro MS/MS in genere priva di segnali che altresì potrebbero falsare la stima dei rapporti delle intensità necessaria per la quantificazione. A seconda del numero dei campioni che vengono esaminati nel corso di un esperimento, in questa porzione dello spettro si può rivelare una coppia, una tripletta o una tetrade di segnali di

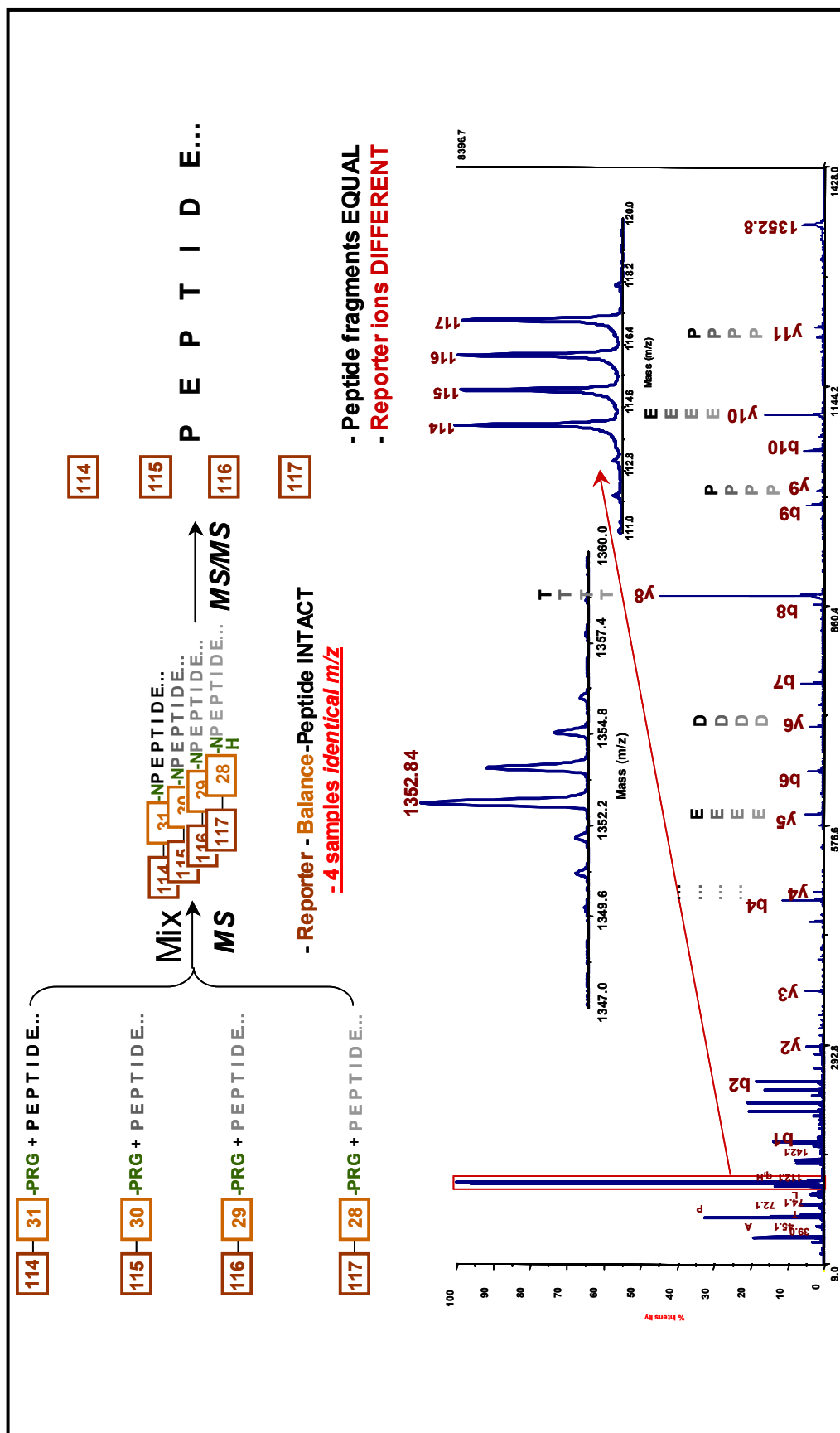


Figura 1.6 Spettri MS ed MS/MS dopo derivatizzazione dei peptidi con il reagente iTRAQ.

intensità differenti. La stima dei rapporti tra questi segnali reporter permette di ricavare l'abbondanza relativa del peptide, quindi conseguentemente della proteina a cui questo appartiene, nel campione di origine.

Le principale limitazione connessa all'uso dell'iTRAQ, origina dalla sua estrema reattività con l'acqua, per cui è assolutamente necessario che i campioni da marcare siano completamente portati a secchezza. L'H₂O non è però il solo composto che ha un'elevata reattività verso l'agente derivatizzante quindi, occorre assicurarsi che i campioni utilizzati per la marcatura siano estremamente puliti. Per lo stesso motivo la metodica non è particolarmente efficiente quando si hanno a disposizione dei campioni provenienti da gel.

1.2.3 Modificoma: proteomica e studio delle modificazioni post-traduzionali (PTMsoica)

La struttura primaria di una molecola proteica e conseguentemente la sua funzionalità possono essere alterate se le proteine sono soggette a delle modifiche post-traduzionali (PTMs). *In primis* un singolo gene può dare origine attraverso il processo di splicing alternativo a molteplici mRNA che sono tradotti in proteine simili ma ben distinte. La creazione del carattere individuale di una proteina prosegue durante il processo di sintesi, secrezione o localizzazione nel suo sito target specifico, tutti step durante i quali la molecola può essere ulteriormente modificata. Alcune PTMs, come ad esempio la fosforilazione o la metilazione, possono essere transienti od avere una natura permanente, altre restano invece immutate una volta apportate come nel caso del processamento proteolitico. Tra tutte le possibili modifiche che oggi si conoscono [30], solo alcune sono state studiate in maniera sistematica [31]. La 2-DE si è ben adattata allo studio delle PTMs perchè permette di avere una visualizzazione delle modifiche e allo stesso tempo di ottenere una stima semi-quantitativa dell'ammontare delle stesse.

Nei processi di mappatura su larga scala, esistono due modi per valutare le PTMs. Il primo metodo, può essere definito come "globale" e prevede la preventiva identificazione delle proteine in base ai peptidi non modificati, seguita dalla ricerca per tali molecole dei possibili peptidi modificati [32]. La seconda metodica, che invece si può indicare come "orientata", focalizza l'attenzione su uno specifico tipo di PTM che può essere a carico di tutte le proteine del campione in esame. Ciò che spesso complica l'analisi relativa alle PTMs è la bassa stechiometria, il basso

fattore di risposta dei peptidi portanti la modificazione o la natura labile della stessa [33]. E' per queste ragioni che le tecniche impiegate si basano per lo più su una selezione d'affinità per la modificazione di interesse. Di certo la PTM maggiormente studiata fino ad ora è stata la fosforilazione per la quale diverse strategie sono state elaborate e largamente utilizzate. Una prima metodica usa anticorpi contro gli amminoacidi fosforilati che permettono di realizzare un'immunoprecipitazione ed il conseguente arricchimento delle proteine modificate a partire da una miscela proteica complessa come può essere un lisato cellulare [34]. Altra strategia è quella che prevede la ritenzione selettiva dei fosopeptidi attraverso "Immobilized Metal Ion Affinity Chromatography" (IMAC), basata sull'alta affinità dei gruppi fosfato per i cationi metallici come il ferro [35, 36]. I fosopeptidi sono catturati selettivamente sulle colonne IMAC a causa dei gruppi fosfato carichi negativamente. Sulla base di questo, un modo per incrementare la specificità della tecnica, risiede nella possibilità di neutralizzare i residui amminoacidici carichi negativamente prima di caricare il campione sulla colonna. L'applicazione di questo tipo di approccio alle cellule di lievito ha consentito l'identificazione di centinaia di fosopeptidi [37]. Un'ulteriore e recente ricerca ha dimostrato le notevoli capacità del bi-ossido di titanio (TiO_2) nelle tecniche di arricchimento dei fosopeptidi prima dell'analisi in massa [38]. L'utilizzo di microcolonne di TiO_2 in cui i peptidi vengono caricati direttamente utilizzando la matrice (DHB) è stato dimostrato riduce drasticamente la percentuale di peptidi non fosforilati che si legano alla fase stazionaria.

Altre modificazioni, come ad esempio le glicosilazioni, iniziano ad essere sistematicamente studiate utilizzando colonne cromatografiche con lectina immobilizzata, in grado di interagire specificatamente con i gruppi glucidici. I campioni arricchiti in glicoproteine, tramite quest'approccio, sono poi digeriti ed i glicopeptidi selezionati attraverso un secondo step di interazione con la lectina. Generalmente, il gruppo saccaridico viene distaccato per procedere alle ulteriori caratterizzazioni utilizzando un'ampia gamma di enzimi. Per i carboidrati N-legati il sito di modificazione può essere determinato attraverso la deglicosilazione in un tampone contenente 50% H_2^{18}O . A seguito di tale processo, l'asparagina inizialmente glicosilata viene trasformata in un residuo di aspartato ed il peptide corrispondente può essere identificato grazie alla presenza nello spettro di massa di un doppietto con una differenza di 2 Da, permettendo una sicura localizzazione della glicosilazione.

E' anche possibile analizzare modificazioni a carico di proteine solo attraverso la spettrometria di massa tandem senza realizzare nessun processo di derivatizzazione o arricchimento. Questo approccio è basato sul fatto che diverse modificazioni danno origine a frammenti ionici marker dopo dissociazione indotta da collisione (collision-induced dissociation CID). I frammenti modificazione-specifici possono essere utilizzati per esperimenti selettivi come "neutral loss" o "precursor scanning".

CAPITOLO 2

SPETTROMETRIA DI MASSA: TECNICA D'ELEZIONE PER L'INDAGINE PROTEOMICA

Uno spettrometro di massa consiste di almeno tre principali componenti (Fig. 2.1):

- a) la Sorgente Ionica
- b) l'Analizzatore di Massa
- c) il Detector

Gli ioni sono prodotti nella sorgente ionica, separati nell'analizzatore in base al rapporto massa/carica (m/z) ed infine vengono rilevati dal detector che genera dei segnali proporzionali all'abbondanza ionica relativa. In generale, una grafico che rappresenta l'abbondanza ionica relativa verso il rapporto m/z è registrato ed analizzato tramite un computer fornendo uno spettro di massa.

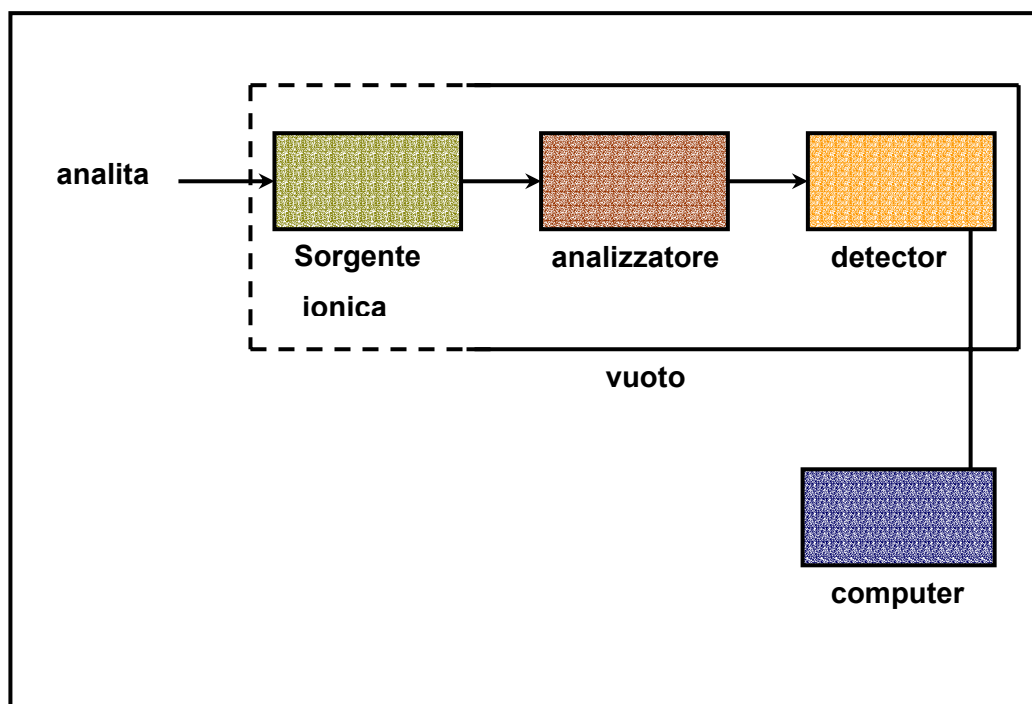


Figura 2.1 Diagramma rappresentante le principali componenti di uno spettrometro di massa: la sorgente ionica, l'analizzatore ed il detector. L'analizzatore ed il detector sono sempre sotto alto vuoto, mentre la sorgente ionica può anche operare a pressione atmosferica.

2.1 Formazione degli ioni in fase gassosa

Sebbene tutti i principi dei moderni analizzatori di massa erano ben noti sin dagli anni quaranta [39], la ionizzazione di biopolimeri intatti come peptidi, proteine, carboidrati ed oligonucleotidi è stata per lungo tempo uno step limitante per l'analisi degli stessi. Fino alla fine degli anni sessanta, la ionizzazione veniva realizzata attraverso impatto elettronico (EI) [40, 41] o ionizzazione chimica (CI) [42]. Nell' EI, la ionizzazione del campione avviene mediante il bombardamento di molecole gassose con un fascio di elettroni ad alta energia. La CI produce ioni attraverso un processo, relativamente blando, di trasferimento di protone. Le molecole di campione vengono esposte ad un forte eccesso di gas reagente ionizzato. Il trasferimento di un protone ad una molecola M del campione da parte del gas produce uno ione positivo $[M+H]^+$. Entrambe queste tecniche di ionizzazione hanno lo svantaggio di richiedere un campione da analizzare relativamente volatile ed inducono un'estensiva decomposizione dell'analita limitando il loro uso a piccole molecole, risultando non adatte alle biomolecole. Tutto questo ha fornito lo stimolo per gli addetti al settore a ricercare dei metodi di ionizzazione alternativi che potessero permettere il superamento di tali limitazioni. Il primo importante passo in questa direzione è stato fatto grazie all'introduzione del desorbimento da campo (Field Desorption, FD) [43], seguito dalla spettrometria di massa di ioni secondari (Secondary Ion Mass Spectrometry, SIMS) [44], bombardamento con atomi veloci (Fast Atom Bombardment, FAB) [45] e desorbimento in plasma (Plasma Desorption, PD) [46, 47]. Seppure queste innovative tecniche di ionizzazione hanno avuto un grande impatto sulla spettrometria di massa applicata alla biologia, rimaneva comunque un limite sulle masse delle molecole che potevano essere ionizzate, non più grandi di 20 kDa. Inoltre, se a questo si aggiunge la necessità di avere a disposizione grandi quantitativi di campione, si capisce perchè nessuna di queste divenne una tecnica di ionizzazione ampiamente diffusa. E' solo con lo sviluppo della Ionizzazione Laser Assistita da Matrice (Matrix Assisted Laser Desorption Ionization, MALDI) [48, 49] e della Ionizzazione per Elettrobulizzazione (Electrospray Ionization, ESI) [50], i cosiddetti metodi di ionizzazione "soft" sviluppati a metà degli anni ottanta, che la spettrometria di massa entrò in una nuova era, quella della spettrometria di massa biologica.

2.1.1 Ionizzazione Laser Assistita da Matrice (MALDI)

La tecnica MALDI (Matrix Assisted Laser Desorption Ionization) deve la sua esistenza alle metodiche sviluppate in precedenza come il FAB e la PD. Michael Barber è stato il primo ad utilizzare una matrice liquida [45] per inglobare l'analita prima della sua ionizzazione attraverso l'irradiazione con atomi di argon o xenon. Il concetto di matrice era anche alla base dei successi della PD quando si introdussero la nitrocellulosa [51] ed il glutatione [52], permettendo la ionizzazione di proteine di 20 kDa. Nel contempo, la ionizzazione laser [53], che era principalmente utilizzata per lo studio dei composti inorganici, iniziava a spostarsi verso una possibile applicazione nella ionizzazione di campioni non volatili [54]. I primi veri e propri successi arrivarono però quando indipendentemente due gruppi, quello di Hillenkamp e Karas (1985-1988) e quello di Tanaka e collaboratori (1988), realizzarono la combinazione della tecnologia laser con l'utilizzo di una matrice per assistere il desorbimento/ionizzazione degli analiti (come per il FAB), consentendo a molecole con masse superiori a 100 000 Da di essere desorbite e misurate. La tecnica andò incontro ad un rapidissimo sviluppo negli anni successivi, determinando il cambiamento della lunghezza d'onda ottimale del laser [55, 56], la scoperta di nuove matrici [56-59] e l'evoluzione dei metodi di preparazione del campione [60, 61], diventando una delle principali metodiche per l'analisi di grandi molecole non volatili.

Le matrici che si usano nel MALDI sono costituite da piccole molecole che permettono di assorbire la radiazione UV/IR. Le prime matrici ad essere usate da Karas ed Hillenkamp erano costituite da triptofano [62] e poi acido nicotinico [63] per le quali veniva utilizzato un laser Nd-YAG quadruplicato in frequenza di 266 nm. Tanaka e colleghi [49] usarono invece glicerolo e polvere di cobalto finemente tritata ed un laser N₂ con una frequenza di 337 nm. Un elevato numero di materiali differenti sono stati testati come matrici e solamente pochi sono risultati essere soddisfacenti. Tra le principali matrici diventate oggi di uso routinario nelle analisi di proteomica (peptidi e proteine) si annoverano l'acido sinapinico, l'acido α -ciano-4-idrossicinnamico (CHCA) [59] ed alcuni derivati benzoici come l'acido di-idrossibenzoico (DHB) [58] (Tabella 2.1). Molte ricerche sono state anche fatte per il miglioramento delle matrici utilizzando il mix di differenti composti come nel caso della super DHB [66, 67] o provando l'aggiunta di additivi quali la nitrocellulosa [68-69]. Il protocollo di preparazione del campione per

MATRICE	Peptidi	Proteine piccole	Proteine grandi	Glicopeptidi	Glicoproteine	Oligonucleotidi
4 HCCA (o CHCA) <i>(acido α-ciano-4-idrossicinnamico)</i>	+++	+	NO	++	NO	NO
SA <i>(acido sinapinico)</i>	+	+++	+++	+	++	NO
DHB <i>(acido 2,5-di-idrossibenzoico)</i>	++	++	+++	+++	++	NO
THAP <i>(2,4,6-tri-idrossiacetofenone)</i>	++	NO	NO	NO	NO	++
HPA <i>(acido 3-idrossipicolinico)</i>	NO	NO	NO	NO	NO	+++
ATT <i>(6-aza-2-tiotimina)</i>	NO	NO	NO	NO	NO	+++
+++ buona; ++ discreta; + sufficiente; NO da evitare.						

Tabella 2.1 Schema delle principali matrici utilizzate nella spettrometria di massa MALDI e relativo indice di compatibilità per l'analisi di differenti biopolimeri [64, 65].

l'analisi MALDI (metodo "dried-roplet") generalmente prevede la deposizione di una soluzione diluita dell'analita di interesse su di una soluzione altamente concentrata di matrice posta su di un supporto (target) metallico. La deposizione delle due soluzioni, può avvenire in successione o allo stesso tempo mescolando i due liquidi che in entrambi i casi vengono lasciati asciugare.

Il processo di desorbimento/ionizzazione dell'analita è stato descritto come una rapida transizione dallo stato solido alla fase gassosa innescata dall'assorbimento dell'energia del laser, seguita da un'espansione a spruzzo della piuma di gas nel vuoto (Fig. 2.2) [64-70].

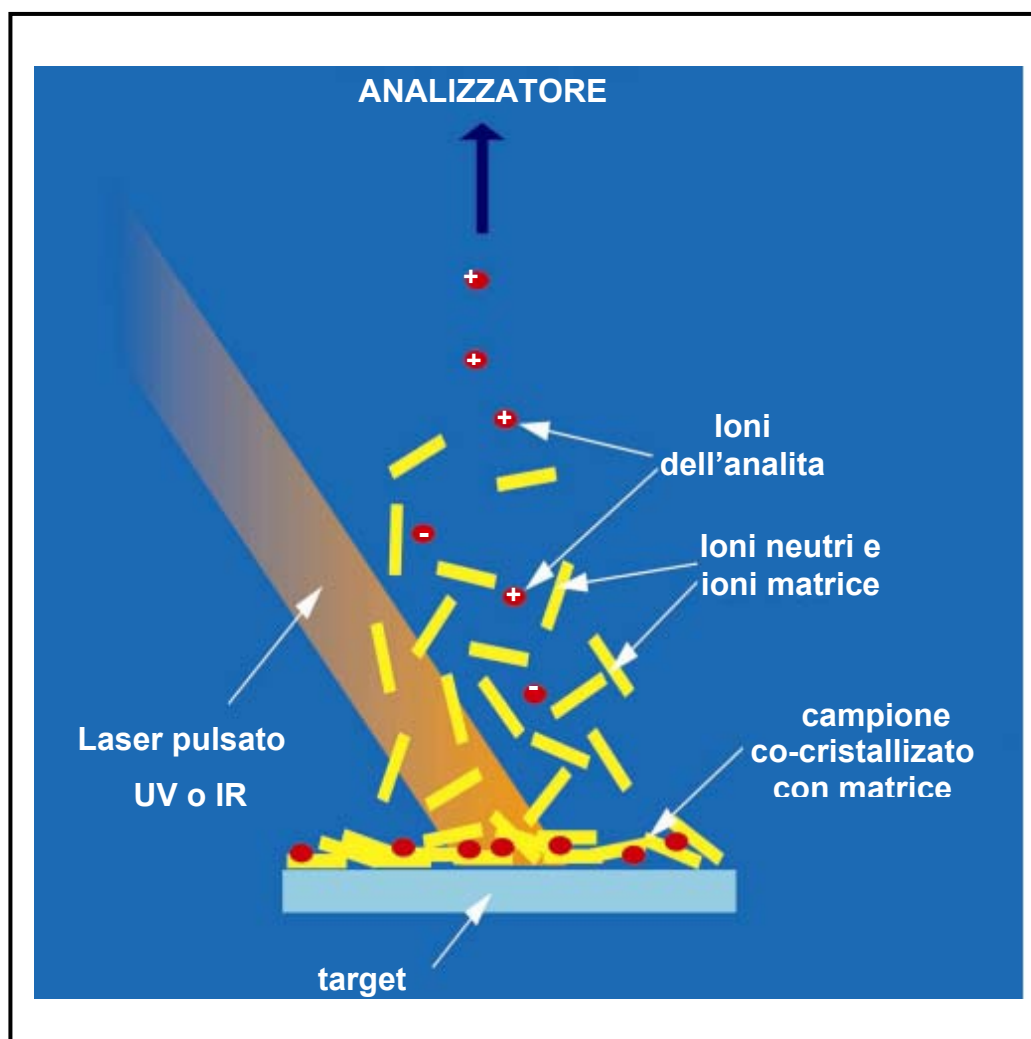


Figura 2.2 Rappresentazione schematica del processo di desorbimento ionizzazione MALDI

Nella piuma in espansione, le molecole dell'analita sono ionizzate attraverso un processo che è ancora per gran parte incompreso e sul quale si

cerca tutt'ora di fare chiarezza [70]. L'energia del laser viene trasferita al campione in un lasso di tempo estremamente breve (tipicamente si hanno ampiezze di pulsazione di pochi ns) così da garantire che sia evitata la decomposizione termica delle molecole. Anche il flusso del laser ha la sua importanza ed è controllato attraverso un sistema di attenuazione del raggio. Esistono delle soglie energetiche che vengono mantenute in maniera da evitare la decomposizione del campione, l'incremento del rumore chimico di fondo e la saturazione del detector [71].

Le molecole della matrice svolgono tre funzioni generali ben definite:

- a) assorbono gran parte dell'energia del laser proteggendo le molecole dell'analita dalla foto-dissociazione
- b) isolano spazialmente le molecole dell'analita evitando la formazione di possibili aggregati
- c) danno origine a reazioni foto-chimiche che garantiscono la ionizzazione del campione

Anche se le conoscenze sul processo di ionizzazione MALDI aumentano di giorno in giorno, la selezione delle matrici ottimali o l'ottimizzazione dei protocolli di preparazione dei campioni sono ancora per lo più procedure empiriche che si basano sulla ripetizione dei tentativi. La selezione di un appropriato metodo di preparazione dell'analita è uno dei punti più critici per l'ottenimento di spettri di massa MALDI di alta qualità. Un metodo ottimale dovrebbe provvedere ad una buona sensibilità, accuratezza di massa e riproducibilità oltre a permettere la rivelazione delle masse ioniche in un ampio range. Altre importanti caratteristiche includono la rapidità della procedura di preparazione e la possibilità di un adattamento all'automatizzazione del processo. Sebbene siano stati testati ripetutamente un elevato numero di nuovi protocolli, il metodo originale del "dried-droplet" proposto da Karas e collaboratori [48] è ancora ampiamente utilizzato in moltissimi laboratori per la sua semplicità, efficienza e la capacità di tollerare moderate concentrazioni di contaminanti. Tra le differenti alternative proposte tre sono quelle che hanno evidenziato la possibilità di ottenere migliori spettri MALDI: la tecnica di evaporazione rapida [61], il metodo "sandwich" [72] e la deposizione del campione per elettro-nebulizzazione [73].

Tra i punti di forza della tecnologia MALDI, oltre a quello di essere un metodo "soft", dobbiamo ricordare: la versatilità nei confronti dei differenti tipi di molecole che possono essere analizzate, l'alta sensibilità, la capacità di consentire l'analisi di sistemi complessi di biomolecole (es. digeriti delle molecole proteiche)

[56], la tolleranza verso concentrazioni relativamente alte di contaminanti (tamponi o sali) [74] ed infine le altre rese [75]. Tutte queste motivazioni, spiegano bene perchè la spettrometria MALDI-TOF è considerata una delle tecniche più potenti negli studi di proteomica.

2.1.2 Ionizzazione per Elettro-nebulizzazione (ESI)

Alla fine degli anni sessanta, un lavoro di tipo pioneristico sulla ionizzazione per elettro-nebulizzazione, volta a generare raggi molecolari di macro-ioni, è stato portato avanti da Dole e collaboratori [76]. Venti anni dopo Fenn e colleghi (1984) [50] svilupparono ulteriormente la tecnica che si mostrò potenzialmente promettente quando lo stesso Fenn presentò i primi spettri di proteine intatte altamente cariche alla Società Americana di Spettrometria di Massa nel 1988 [77]. Oggi questa tecnica, rappresenta un mezzo indispensabile in un ampio campo di applicazioni utilizzando la spettrometria di massa e nello studio della scienza delle proteine.

La ionizzazione elettrospray, comporta la nebulizzazione della soluzione di analita (disciolto in una miscela di solventi polari come acqua e metanolo), attraverso un capillare di conduzione (velocità di flusso: 3-20 $\mu\text{l}/\text{min.}$), mediata dall'applicazione di un campo elettrico tra la sua punta e l'orifizio dello spettrometro di massa. La soluzione, che giunge sulla punta del capillare, è attratta verso il contro-elettrodo carico in maniera opposta mentre la tensione superficiale agisce in senso inverso. Quando la forza elettrostatica del voltaggio applicato vince la tensione superficiale della soluzione di campione, uno strettissimo cono di liquido (cono di Taylor) si forma sulla punta del capillare e si generano delle piccole goccioline cariche che vanno a creare uno spruzzo (Fig 2.3). Mentre le gocce si spostano verso l'orifizio dello spettrometro di massa, evaporano e si riducono di dimensioni fino a raggiungere un punto, definito come limite di Rayleigh, in cui la forza columbiana di superficie supera la tensione superficiale determinando l'esplosione della goccia e la formazione di nuove goccioline dalle dimensioni sempre più ridotte (Fig 2.3) [78]. Questo processo continua fino al punto in cui uno ione desorbe dalla superficie della gocciolina (modello di evaporazione dello ione)

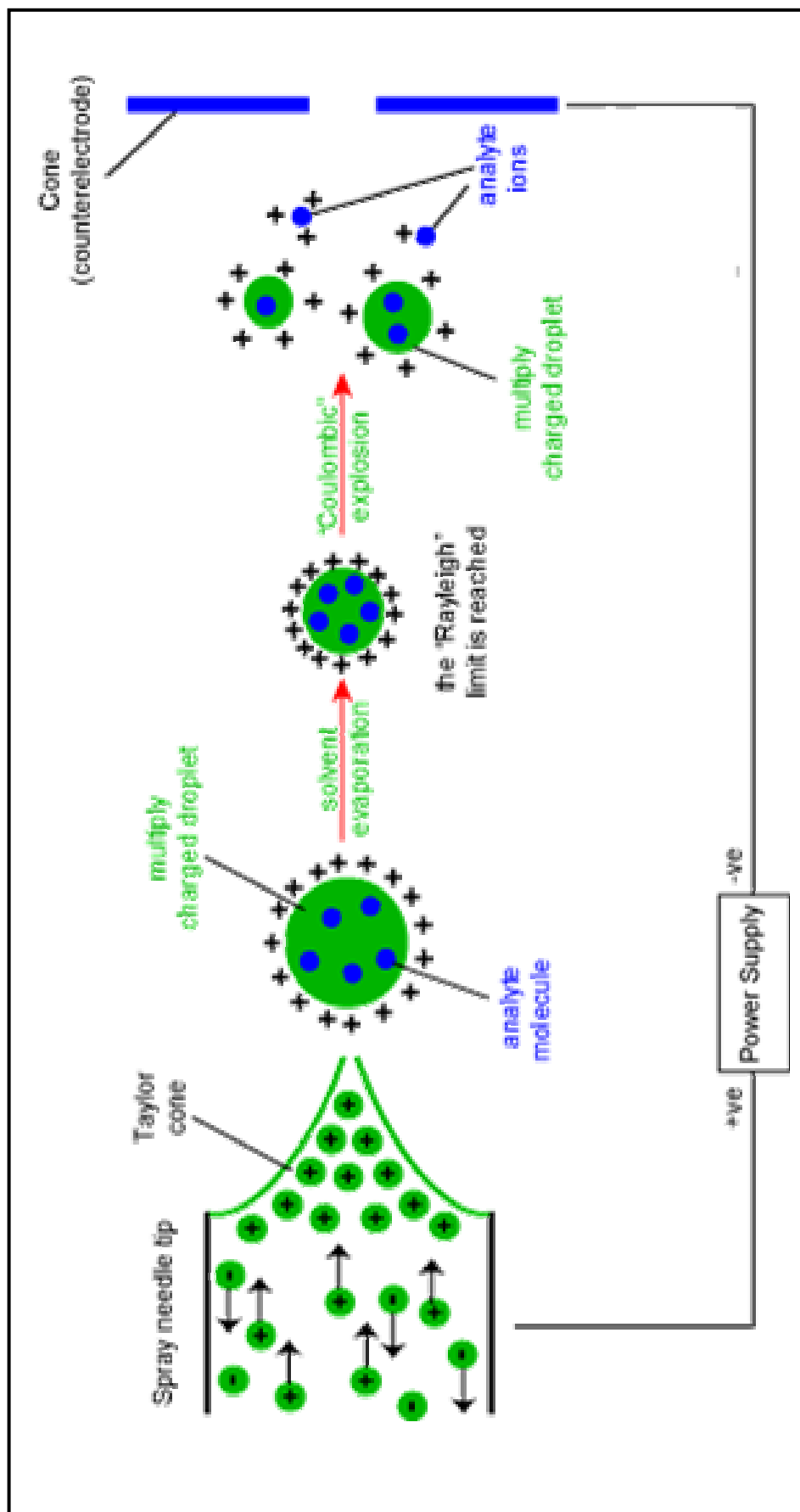


Figura 2.3 Rappresentazione schematica delle varie fasi del processo di ionizzazione elettrospray. Adattato da [82].

[79, 80] o il solvente è completamente rimosso (modello della carica residua) [76, 81]. Il meccanismo esatto della formazione degli ioni, se sia per evaporazione degli ioni stessi o per completa rimozione del solvente, è ancora ampiamente discusso, e probabilmente meccanismi differenti possono realizzarsi in differenti situazioni.

Il processo di elettronebulizzazione è dipendente dalla concentrazione piuttosto che dalla massa ed incrementi di sensibilità si ottengono aumentando la concentrazione dell'analita e riducendo i volumi di campione [83]. Da questi presupposti, si è partiti per elaborare migliorie a carico della sorgente di ionizzazione elettrospray. I primi risultati sono stati ottenuti da Gale e Smith [84] a cui hanno fatto seguito Emmett e Caprioli con il loro "micro-elettrospray" [85] riducendo i flussi a 300-800 nl/min, Wilm e Mann con la ionizzazione "nano-elettrospray" [86, 87] ed infine Tempst e collaboratori che con la loro "injection adaptable fine ionisation source" (JaFIS) [88] si sono spinti sino a flussi di 1-2 nl/min.

2.2 Analizzatore

2.2.1 Analizzatore a tempo di volo (TOF)

Il concetto alla base dell'analizzatore a tempo di volo (Time of Flight, TOF) risale alla metà degli anni quaranta quando Stephens suggerì che uno spettrometro di massa poteva essere realizzato con un campo elettrico in grado di separare i pacchetti di ioni sulla base del diverso rapporto m/z [89]. Due anni più tardi, il primo strumento TOF dotato di sorgente pulsata, una regione di deriva libera da campo, ed un sistema di rivelazione fu realizzato da Cameron ed Eggers (1948) [90]. Ad ogni modo, questi strumenti non suscitarono nessun interesse a causa della bassa risoluzione che garantivano e delle imperfezioni dell'elettronica. Gli analizzatori a tempo di volo dovettero aspettare la nascita di altre innovazioni, come la focalizzazione ritardata (time lag focusing) [91], l'estrazione ionica ritardata (delayed ion extraction) [92-94], specchi ionici (ion mirrors) [95] e le sorgenti MALDI, prima che i loro principali parametri analitici raggiungessero una risoluzione di oltre 20 000 (full width half maximum, FWHM) ed un'accuratezza di massa al di sotto dei 50 ppm. Tutto questo ha permesso agli analizzatori TOF di poter essere considerati uno dei mezzi più importanti per la ricerca in campo biologico [96-98].

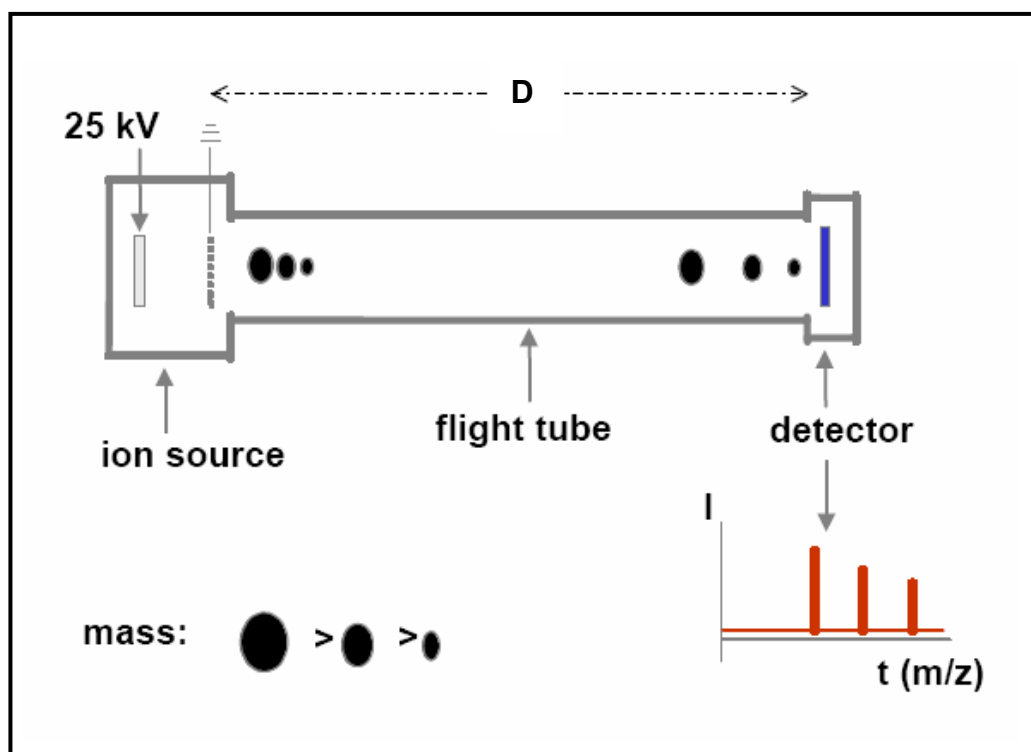


Figura 2.4 Rappresentazione schematica di uno spettrometro di massa equipaggiato con un analizzatore a tempo di volo (TOF) in modalità lineare

2.2.1.1 Principi dell'analizzatore TOF

La semplicità dei principi di base dell'analizzatore TOF può dare in parte una spiegazione del suo attuale successo. Negli spettrometri TOF che lavorano in modalità "lineare" (Fig. 2.4), gli ioni generati in sorgente sono accelerati, attraverso l'applicazione di un campo elettrico (E), lungo una regione di deriva libera dal campo (tubo di volo, D). Idealmente gli ioni entrano nel tubo di volo con la stessa energia cinetica (E_k),

$$E_k = zeE$$

(dove z è il numero delle cariche ed e la carica di un elettrone) ed

$$E_k = 1/2mv^2$$

Le velocità degli ioni (v) sono però dipendenti dalla loro massa (m),

$$v = (2zeE/m)^{1/2}$$

Il tempo (t) necessario per attraversare la regione di deriva è quindi registrato e convertito in uno spettro di massa,

$$D = vt$$

(con D pari alla distanza coperta dallo ione)

$$t = (m/2zeE)^{1/2}D \rightarrow \boxed{m/z = 2eE(t/D)^2}$$

2.2.1.2 Evoluzione dell'analizzatore a tempo di volo

Seppure i principi alla base degli analizzatori TOF erano molto semplici, gli strumenti equipaggiati con tubi di volo non raggiunsero da subito ed in maniera diretta il successo per due motivazioni fondamentali. In primo luogo mostravano una bassa risoluzione ed accuratezza di massa ed inoltre erano difficilmente interfacciabili con sorgenti ioniche continue. I punti più critici erano costituiti dalle conseguenze delle condizioni iniziali di formazione degli ioni (temporali, spaziali e dalla distribuzione delle energie cinetiche), dal loro destino durante il processo di frammentazione (frammentazione metastabile) ed infine dalle proprietà del sistema di registrazione (velocità di campionamento, ampiezza di banda e risposta del detector).

Wiley e McLaren [91] furono i primi a descrivere gli effetti delle variabili che influenzavano la formazione degli ioni, la disposizione nel campo di estrazione e l'energia cinetica sulla risoluzione delle masse. Loro stessi svilupparono la teoria del "time lag focusing" per correggere la distribuzione delle energie attraverso l'estrazione ritardata degli ioni dalla sorgente. Gli ioni sono generati in una regione priva di campo prima che a questi venga applicato un impulso di estrazione, ciò genera un ritardo tra la formazione delle specie ioniche e la loro estrazione. Durante questo tempo, gli ioni sono distribuiti nella sorgente sulla base delle loro velocità iniziali e quelli con velocità più elevata riceveranno una minore energia all'applicazione del campo di estrazione consentendo a quelli che invece sono più rallentati di raggiungerli. La distribuzione spaziale fu invece corretta attraverso l'uso di due regioni di estrazione che, attraverso aggiustamenti della loro lunghezza e dell'intensità relativa del campo applicato, hanno permesso la focalizzazione del pacchetto di ioni in corrispondenza del detector. Un ulteriore accorgimento usato per correggere la dispersione energetica si è ottenuto grazie all'impiego di specchi ionici. La tecnica fu elaborata da Mamyrin e collaboratori negli anni settanta [95] e loro stessi definirono il sistema inventato con il termine di

“mass reflectron”. Il reflectron (Fig. 2.5) è costituito da una serie di anelli e/o griglie a voltaggi crescenti (in modalità positiva) ed è posizionato alla fine della regione di deriva priva di campo. Gli ioni, penetrano il reflectron finché non raggiungono energia pari a zero, tornano indietro e sono riaccelerati attraverso il reflectron stesso dal quale escono con un'energia pari alla loro energia di ingresso. Si ottiene in questo modo una risoluzione ottimale quando il tempo aggiuntivo speso dagli ioni più veloci nel reflectron è pari al tempo aggiuntivo speso dagli ioni più lenti nella regione di accelerazione primaria e nel tubo di volo.

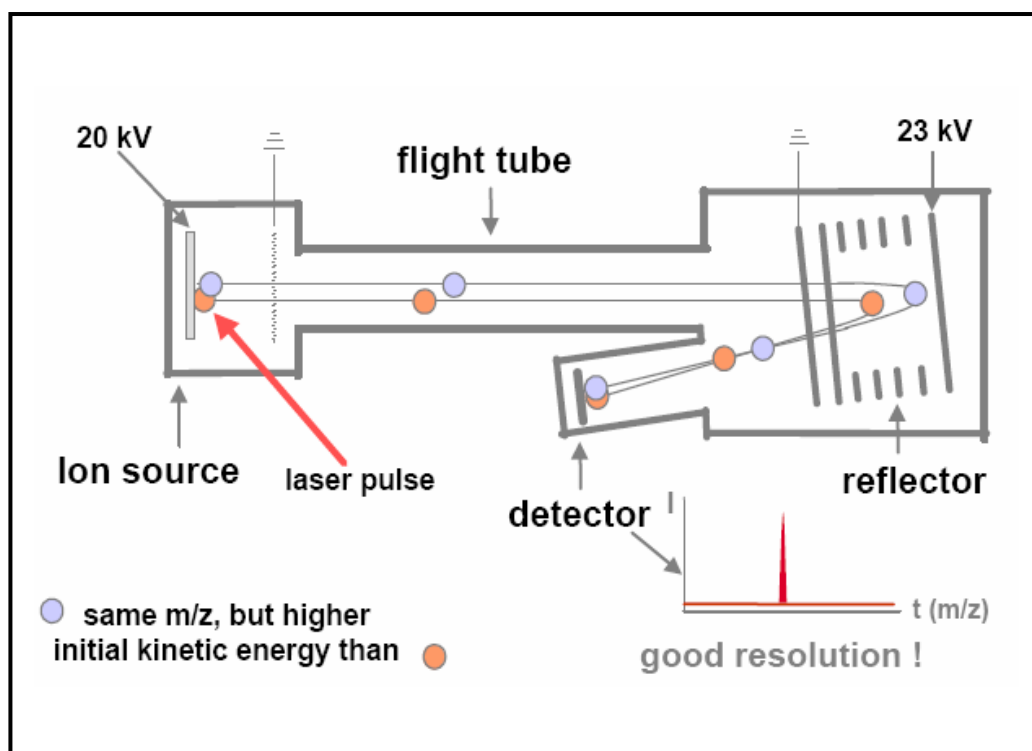


Figura 2.5 Rappresentazione schematica di uno spettrometro di massa equipaggiato con un analizzatore a tempo di volo (TOF) in modalità reflectron

Lo sviluppo del tubo di volo ad accelerazione ortogonale (orthogonal acceleration Time-of-Flight, oaTOF) è stato uno dei successivi step di notevole importanza nello sviluppo della tecnologia TOF. Il principio di base del funzionamento dell'oaTOF risale agli anni sessanta e fu sviluppato da O'Halloran ma non suscitò un gran successo fino alla sua riscoperta ad opera di Dawson e Guilhaus nel 1989. Questo approccio fu utilizzato la prima volta per compensare la distribuzione temporale degli ioni quando il TOF venne usato con una sorgente EI.

Gli ioni venivano collimati da una lente elettrostatica e quindi trasferiti in una regione di immagazzinamento. In seguito venivano estratti ed accelerati, in una direzione ortogonale rispetto a quella di formazione, all'interno di un analizzatore TOF lineare. Dodonov e collaboratori nel 1991 implementarono tale tecnologia interfacciando una sorgente ESI con un analizzatore TOF mediante l'utilizzo di un quadrupolo elettrostatico per collimare il fascio ionico così da incrementare la trasmissione e l'accelerazione ortogonale degli ioni verso un analizzatore di massa di tipo reflectron. Tutto questo ha permesso di disaccoppiare il TOF dalle condizioni iniziali di formazione degli ioni, permettendo a questo analizzatore di poter essere interfacciato con qualunque tipo di sorgente [99].

Gli ulteriori sviluppi della tecnologia degli analizzatori a tempo di volo ha portato alla creazione degli spettrometri di massa TOF/TOF [100] che permettono di frammentare un determinato peptide di interesse, rilevarne i frammenti ed ottenere da questi informazioni di sequenza. In questo tipo di strumenti (Fig. 2.6), i due tubi di volo posti in serie ed intervallati dalla presenza di una cella di collisione svolgono funzioni differenti. Il primo TOF permette di selezionare lo ione precursore prescelto che raggiunge la cella di collisione dove generalmente viene frammentato grazie all'interazione con le molecole di un gas inerte (collision induced dissociation, CID). Il secondo tubo di volo permette invece di fare una scansione dei frammenti generati dalla rottura del precursore che, giungendo al rivelatore, forniscono le informazioni per ottenere uno spettro MS/MS.

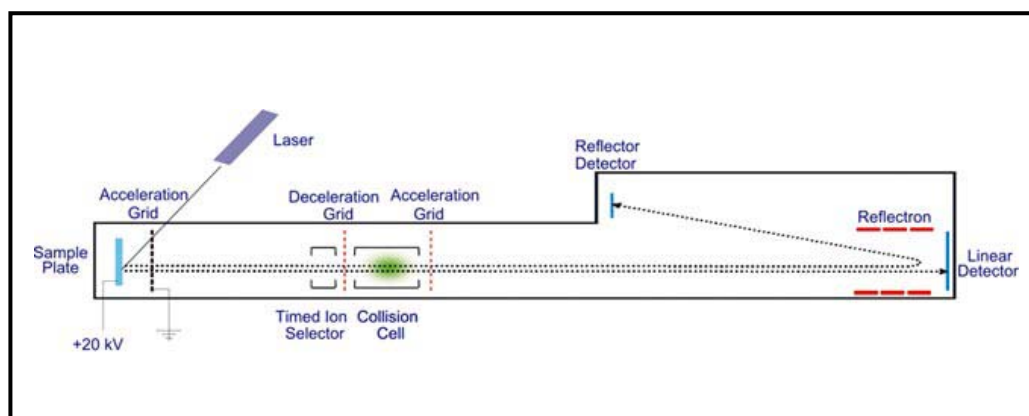


Figura 2.6 Rappresentazione schematica dello spettrometro di massa “4700 Proteomics Analyser” (Applied Biosystems) dotato di analizzatore TOF/TOF.

2.2.2 Analizzatore a quadrupolo (Q)

Gli analizzatori a quadrupolo furono sviluppati da Paul e Raether nel 1955 [101] e costituiscono oggi la gamma di analizzatori più economici distribuiti commercialmente. Il quadrupolo opera utilizzando un campo elettrico dinamico che consente solo agli ioni con un determinato valore del rapporto m/z di essere trasmessi attraverso di esso, ne consegue che per ottenere uno spettro di massa occorre operare facendo una scansione in maniera che differenti ioni possano essere trasmessi sequenzialmente. Questi dispositivi richiedono quindi un continuo flusso ionico rendendoli ottimali per l'accoppiamento con la sorgente di ionizzazione ESI. Un analizzatore quadrupolare è costituito da quattro barre metalliche (generalmente di molibdeno), allineate parallelamente lungo l'asse z . L'analizzatore ideale dovrebbe possedere delle barre di tipo iperbolico consentendo la creazione di un campo lineare e simmetrico [102] (Fig. 2.7), ma per convenienza la maggiorparte è costituita da delle barre cilindriche (Fig. 2.8).

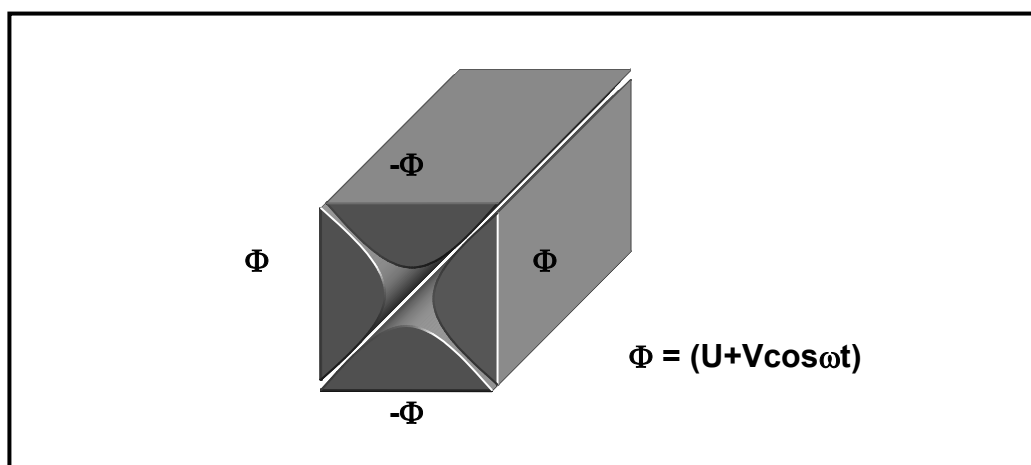


Figura 2.7 Rappresentazione schematica della forma ideale di un quadrupolo.

Ad ogni coppia di barre elettricamente connesse è applicato un voltaggio: U nel caso di un campo elettrico continuo (DC) e V nel caso di un campo elettrico alternato a radiofrequenza (RF). Gli ioni, che uscendo dalla sorgente fanno il loro ingresso nel quadrupolo, risentono delle forze di accelerazione nelle direzioni x ed y indotte dal campo elettrico combinato. Il movimento di uno ione attraverso questo campo è descritto dall'equazione di Mathieu, che definisce regioni di stabilità e di instabilità per gli ioni in un campo elettrico generato da un quadrupolo.

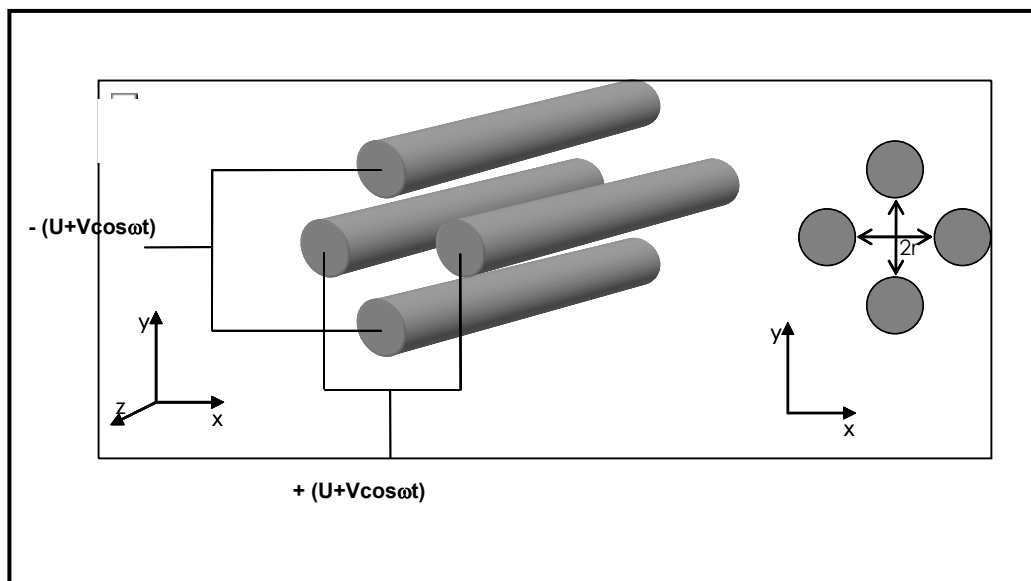


Figura 2.8 Struttura di un analizzatore di massa quadrupolare convenzionale.

Schematicamente, questa equazione si traduce in aree di stabilità che sono funzioni di U e V per ioni con differente massa (Fig. 2.9). Attraverso la variazione dei parametri applicati all'analizzatore, mantenendo però costante il rapporto U/V , è possibile stabilizzare e quindi consentire ad uno ione con specifico valore di m/z di attraversare il quadrupolo. Realizzando una scansione sull'intero range di massa si può ottenere la separazione e la successiva rivelazione di masse differenti a partire da 4 000 Da con una risoluzione di circa 3 000 (FWHM).

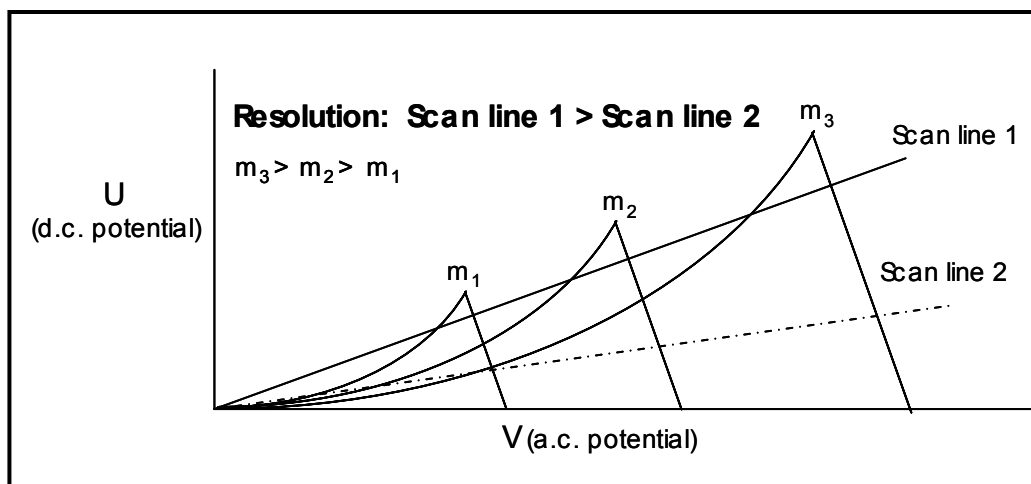


Figura 2.9 Diagramma di stabilità per tre ioni con differente rapporto di massa su carica.

Un quadrupolo, funziona anche in maniera eccellente come cella di collisione per il CID a “bassa” energia (nell’ordine degli eV). L’analizzatore focalizza la traiettoria degli ioni verso il suo centro (asse z), per un effetto collaterale del campo elettrico oscillante. Il processo di CID provoca la dispersione degli ioni e la loro attivazione ed il quadrupolo aiuta la rifocalizzazione degli stessi incrementandone l’efficienza di trasmissione.

2.2.3 Spettrometria di massa tandem (MS/MS)

La spettrometria di massa convenzionale, che fa uso delle sorgenti di ionizzazione “soft” come l’elettrospray o il MALDI, fornisce poche informazioni di tipo strutturale e nella maggiorparte dei casi permette solo l’analisi del peso molecolare dell’analita. Le informazioni sulla struttura del campione possono invece essere ottenute grazie alla spettrometria di massa tandem (MS/MS) che prevede uno step di attivazione/dissociazione. La spettrometria di massa tandem consiste nell’isolare uno ione precursore (ione parente) di interesse, selezionato durante un esperimento di MS, attivare quello ione determinandone la sua frammentazione, ed acquisire lo spettro di massa dei frammenti ionici. Si possono realizzare due diversi tipi di esperimenti di MS/MS. Nel primo caso si può realizzare un tandem “nello spazio”, utilizzando due analizzatori di massa disposti in serie separati da una cella di collisione come negli strumenti a triplo quadrupolo, nei Q-TOF o in quelli TOF/TOF (sezione 1.3.2). Alternativamente si può avere un tandem “nel tempo”, quando invece si fa uso di trappole ioniche (ION TRAP) o di spettrometri di massa a risonanza ionica ciclotronica in trasformata di Fourier (FT-ICR) che permettono di compiere tutti i differenti step, in tempi diversi, utilizzando lo stesso analizzatore.

2.2.3.1 Dissociazione indotta/attivata da collisione (CID/CAD)

La dissociazione indotta/attivata da collisione (Collision Iduced/Activated Dissociation CID/CAD) è la tecnica di attivazione più diffusa in virtù della sua semplicità di uso e del livello di controllo. Per prima venne applicata nei laboratori di Jennings [103] e McLafferty [104] alla fine degli anni sessanta. In generale, la dissociazione procede attraverso due step discreti. In un primo momento gli ioni di interesse vanno incontro ad un aumento della loro energia interna a causa dell’urto anelastico con le molecole o gli atomi di un gas inerte e parte della loro energia

identificati come X, Y e Z. Un numero viene aggiunto come pedice alla lettera corrispondente per indicare dove avviene la frammentazione lungo la catena amminoacidica. Inoltre se la massa differisce di una o due unità di massa (che si assume siano idrogeni) dal frammento prodotto dalla semplice rottura si aggiunge la notazione primo (') o secondo (''). Gli ioni immonio degli amminoacidi sono indicati con la lettera maiuscola relativa al loro nome (es. G per glicina). Questa notazione è stata proposta quando gli studi sulla frammentazione peptidica erano ancora agli inizi. La nomenclatura è stata modificata poi nel 1988 da Biemann [106] e i cambiamenti includono:

- passaggio dal maiuscolo al minuscolo per indicare i frammenti ionici prodotti dalla rottura del legame ammidico (es. A, B e C sono ora indicati come a, b e c). Questo per evitare possibili confusioni con i caratteri usati per gli ioni immonio
- aggiunta delle lettere d, w e v per indicare gli ioni (derivanti da rotture concomitanti dello scheletro principale e delle catene laterali) osservati in esperimenti CID ad "alta" energia

Le strutture dei frammenti ionici e le loro assegnazioni sono mostrate nelle Fig. 2.10, e la loro origine rispetto al peptide in Fig. 2.11 e Fig. 2.12

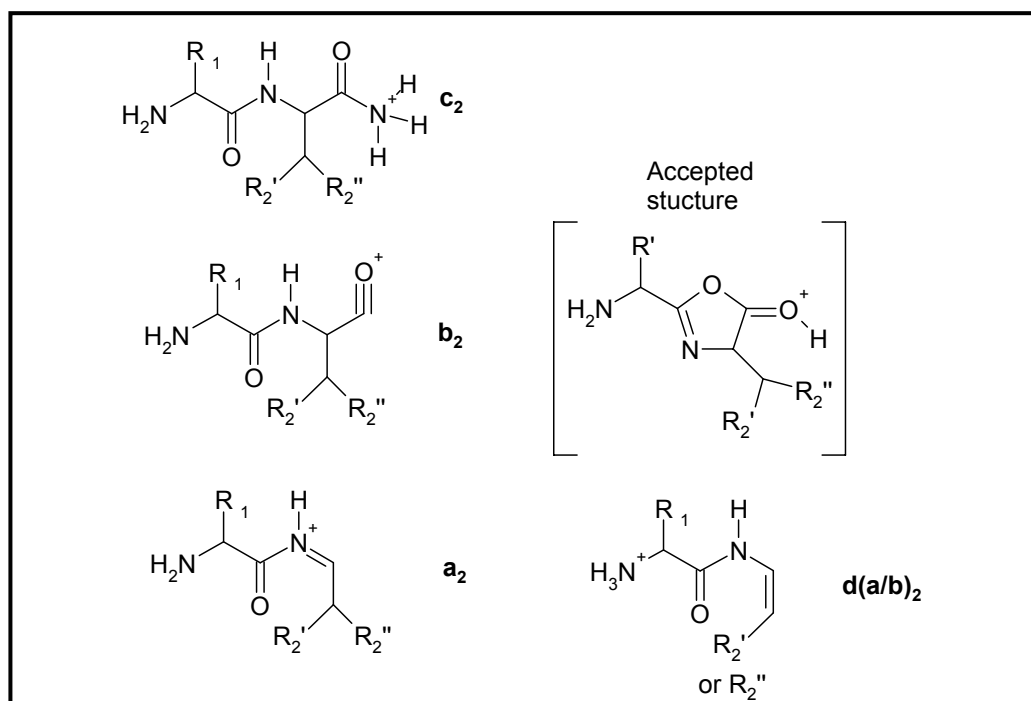


Figura 2.11 Frammenti N-terminali.

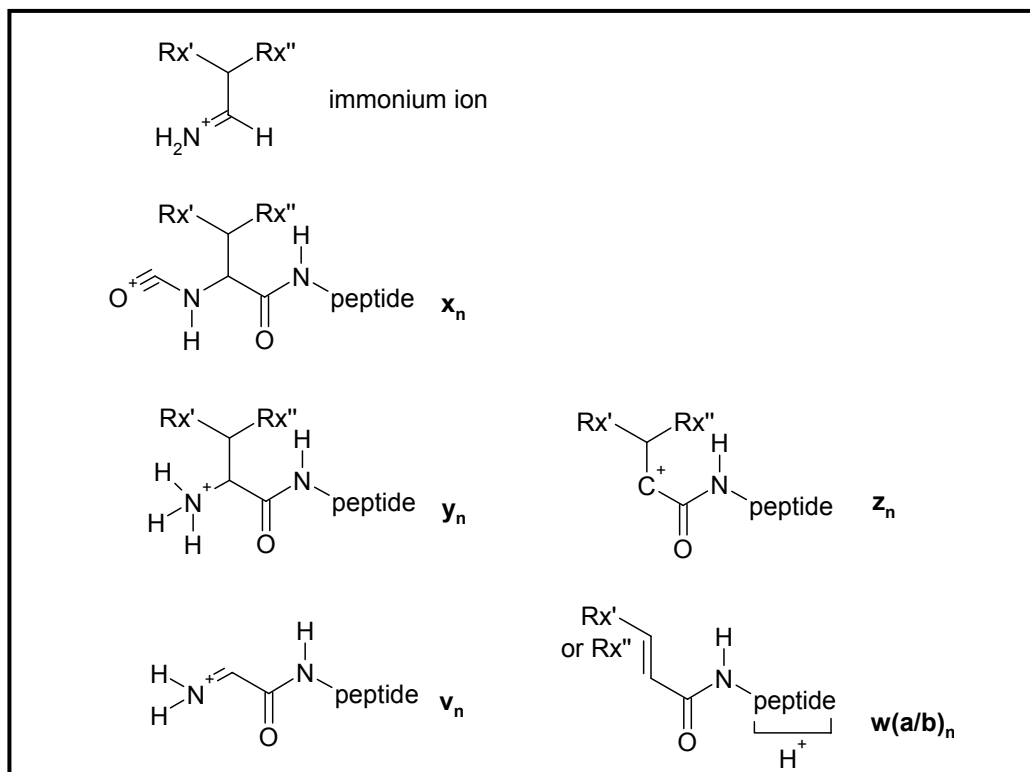


Figura 2.12 Frammenti C-terminali.

2.2.4 Spettrometri di massa ibridi quadrupolo TOF (Q-TOF)

Negli ultimi anni gli spettrometri di massa a triplo quadrupolo, in cui si aveva un quadrupolo che permetteva di compiere la scansione ed ottenere lo spettro di massa, seguito da un secondo che assolveva alle funzioni di cella di collisione e da un terzo che separava i frammenti ottenuti, sono stati ampiamente sostituiti dagli strumenti Q-TOF [107]. Questo tipo di strumenti si basano sull'accoppiamento tra un filtro di massa quadrupolare, una cella di collisione esapolare e un oaTOF (Fig. 2.13). Il Q-TOF consente di ottenere una risoluzione pari a 10 000 ed un'accuratezza di massa di pochi ppm. Le alte performance permettono quindi di ottenere una rapida ed inequivocabile identificazione delle proteine di interesse basata sull'individuazione di uno o pochi peptidi che vengono frammentati. L'eccellente sensibilità, velocità, risoluzione, accuratezza e la stabilità della calibrazione dell' oaTOF evidenzia notevoli vantaggi quando si paragonano i risultati di esperimenti tandem nell'analisi dei prodotti di CID, condotti con questo spettrometro, con quelli degli strumenti a quattro settori.

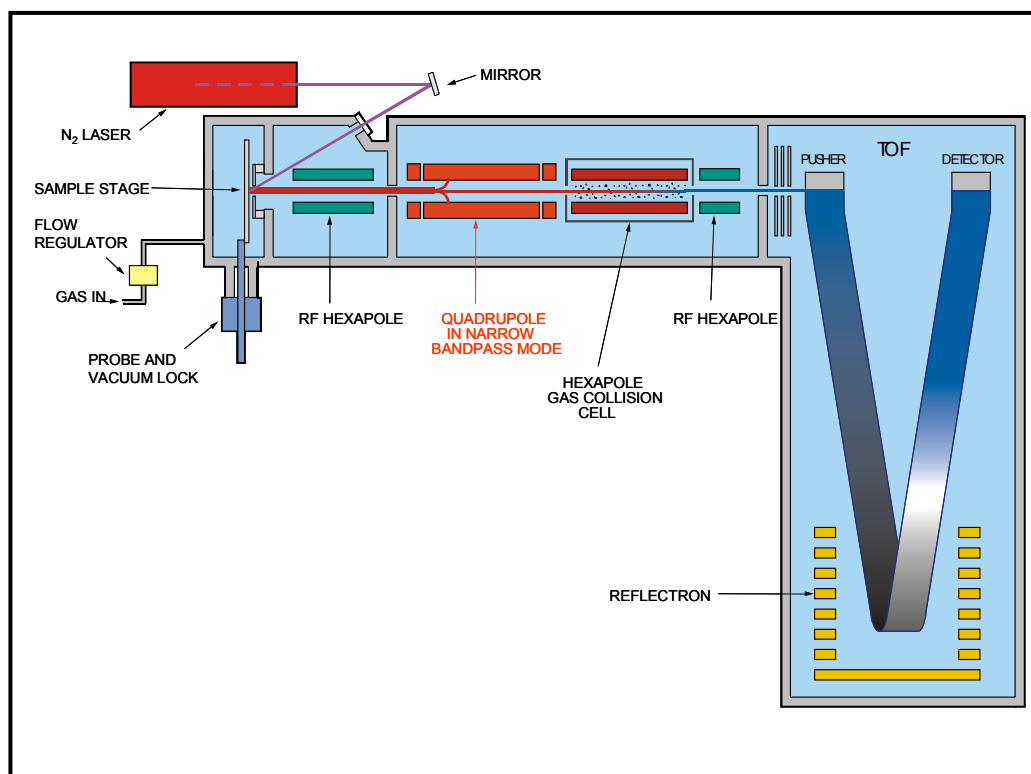


Figura 2.13 Rappresentazione schematica del MALDI-Q-TOF della Micromass in modalità MS/MS.

2.2.5 Analizzatori di massa ad intrappolamento ionico

Gli strumenti che permettono di realizzare esperimenti MS/MS “nel tempo” operano utilizzando campi magnetici e/o elettrostatici statici e/o dinamici. I due strumenti più comuni ad intrappolamento ionico sono costituiti dalla trappola ionica quadrupolare (ION TRAP) [108] e dallo spettrometro di massa a risonanza ionica ciclotronica in trasformata di Fourier (FT-ICR). [109, 110].

Il modo di operare di un' ION TRAP è simile, in qualche modo, a quello di un analizzatore a quadrupolo (sezione 1.3.3), eccetto per il fatto che la trappola utilizza un campo elettrostatico dinamico in tre dimensioni (x, y e z) mentre il quadrupolo opera in un campo dinamico a due dimensioni (x ed y) mentre gli ioni si muovono lungo l'asse z. Gli ioni sono isolati, eccitati e rivelati utilizzando le proprietà della loro stabilità e la manipolazione del campo elettrostatico. Una trappola ionica (Fig. 2.14) consiste di un elettrodo anulare (ring electrode) posto tra due elettrodi a “coperchio” (endcap electrodes). La forma della superficie interna di questi tre elettrodi segue un profilo tridimensionale quasi iperbolico. Gli elettrodi endcap sono dotati di piccoli orifizi che consentono l'ingresso e l'uscita

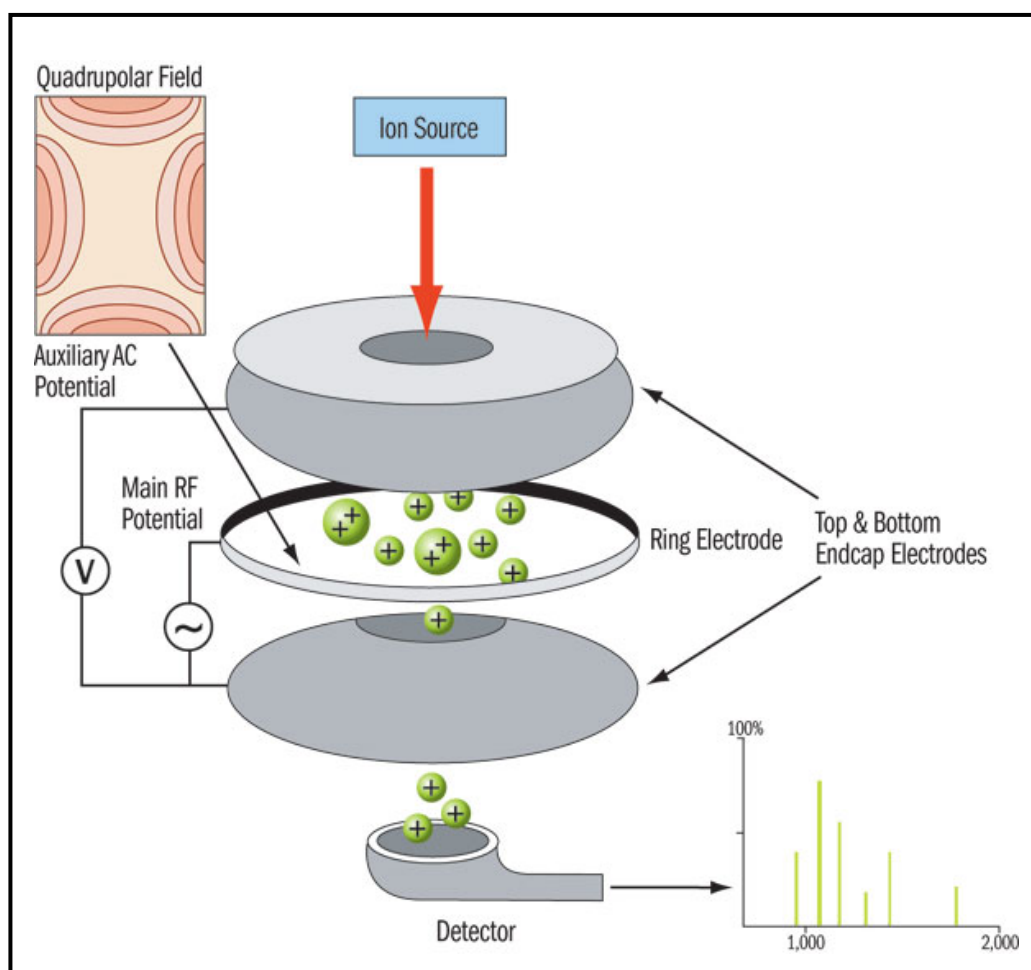


Figura 2.14 Rappresentazione schematica di un analizzatore a trappola ionica (ION TRAP).

degli ioni dalla trappola. Un potenziale RF è applicato all'elettrodo anulare mentre gli elettrodi a "coperchio" sono mantenuti a terra, e tale disposizione permette di generare un campo quadrupolare. In dipendenza dal livello del voltaggio RF, il campo può permettere di intrappolare ioni in un particolare range di massa. Il range delle masse ioniche che possono essere intrappolate simultaneamente nell'analizzatore è descritto dal diagramma di stabilità (Fig. 2.15). Il diagramma di stabilità è un grafico che mostra in corrispondenza di quale particolare potenziale gli ioni con un determinato valore di m/z sono stabili o instabili nel campo. Lo spettro di massa viene prodotto variando il potenziale RF in modo da espellere sequenzialmente dalla trappola verso il rivelatore gli ioni secondo un ordine di rapporto m/z crescente.

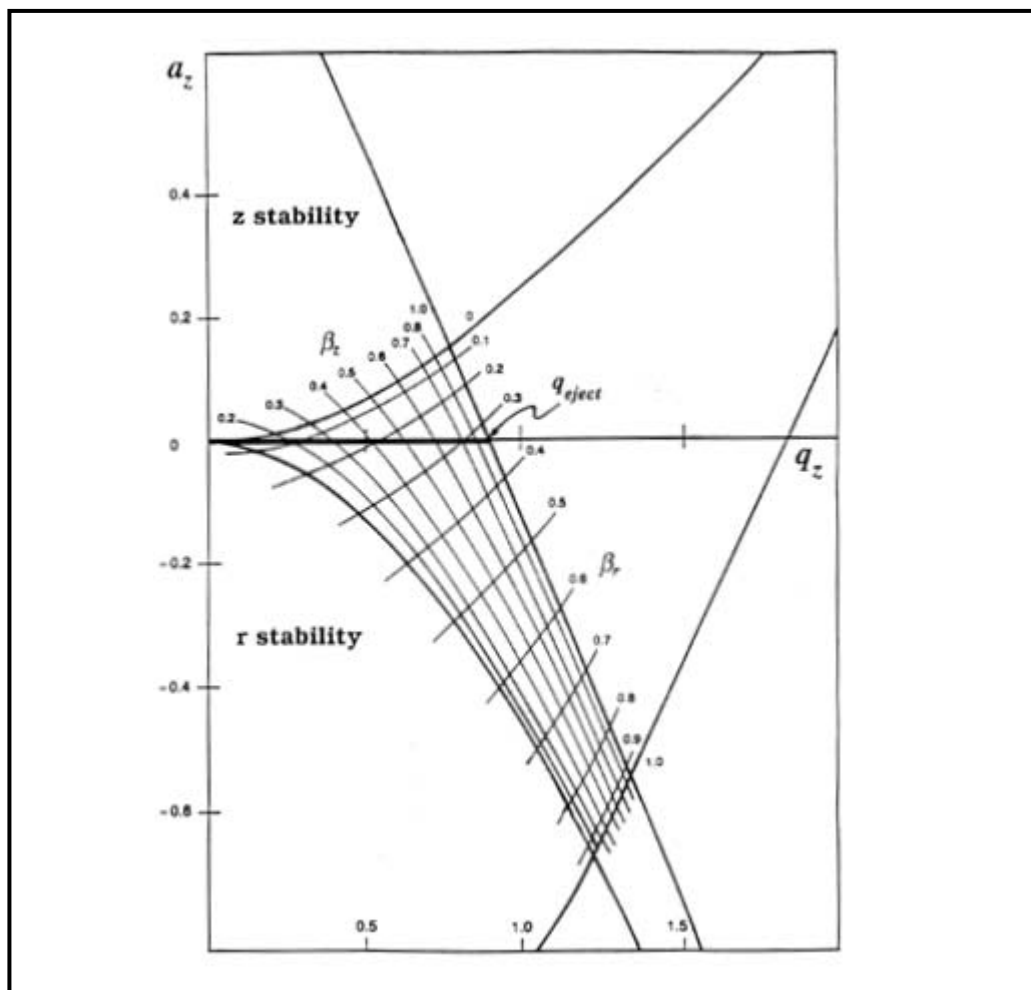


Figura 2.15 Diagramma di stabilità per un analizzatore a trappola ionica.

Gli strumenti FT-ICR sono in grado di intrappolare gli ioni utilizzando un campo magnetico statico, prodotto da un magnete superconduttore. In un campo di questo tipo uno ione sarà forzato a mantenere un moto circolare con una caratteristica frequenza orbitale dipendente dallo specifico rapporto massa su carica. Negli esperimenti di MS/MS gli ioni possono essere eccitati sia attraverso risonanza, collisione con un gas inerte (CID), esposizione a fotoni (infra-red multiphoton dissociation, IRMPD) o a elettroni (electron capture dissociation, ECD). In un analizzatore FT-ICR gli ioni vengono intrappolati in una cella cubica grazie all'applicazione di un impulso a radiofrequenza tra due piastre della cella che fungono da trasmettitore (piastre di eccitazione) producendo un moto orbitale coerente. Gli ioni in moto orbitale generano un debole segnale "immagine" che viene raccolto dalle piastre della cella che agiscono come ricevitore (piastre di rivelazione). La frequenza del segnale generato da un particolare ione è uguale

alla sua frequenza orbitale che a sua volta è inversamente correlata al suo rapporto m/z . L'intensità del segnale per una data frequenza è proporzionale al numero di ioni aventi il corrispondente valore di m/z . Il segnale totale viene amplificato e quindi scomposto nelle varie componenti di frequenza generando così lo spettro di massa. Se la pressione all'interno della cella è molto bassa, ossia se viene mantenuto un vuoto molto spinto, il moto orbitale degli ioni può essere mantenuto per tempi relativamente lunghi e la frequenza misurata ripetutamente numerose volte per ottenere un'accuratezza molto elevata. Gli strumenti FT-ICR sono in grado quindi di produrre spettri di massa con un'altissima risoluzione.

Un grande vantaggio degli strumenti che sono dotati di analizzatori ad intrappolamento ionico risiede nel fatto che questi consentono di realizzare esperimenti di massa tandem a diversi step, non solo MS/MS ma anche MS/MS/MS e più in generale MS^n .

CAPITOLO 3

IDENTIFICAZIONE DELLE PROTEINE

Tradizionalmente, le proteine venivano identificate attraverso esperimenti di *de novo* sequencing utilizzato il procedimento automatizzato conosciuto con il nome di degradazione di Edman. Oggigiorno, questa tecnica è considerata in parte obsoleta ed è stata ampiamente sostituita dalla spettrometria di massa per la sua sensibilità e rapidità. Esistono due tipi differenti di approcci basati sull'utilizzo dell'MS per l'identificazione proteica che vengono definiti rispettivamente con i termini "top-down" e "bottom-up" [111] (Fig. 3.1).

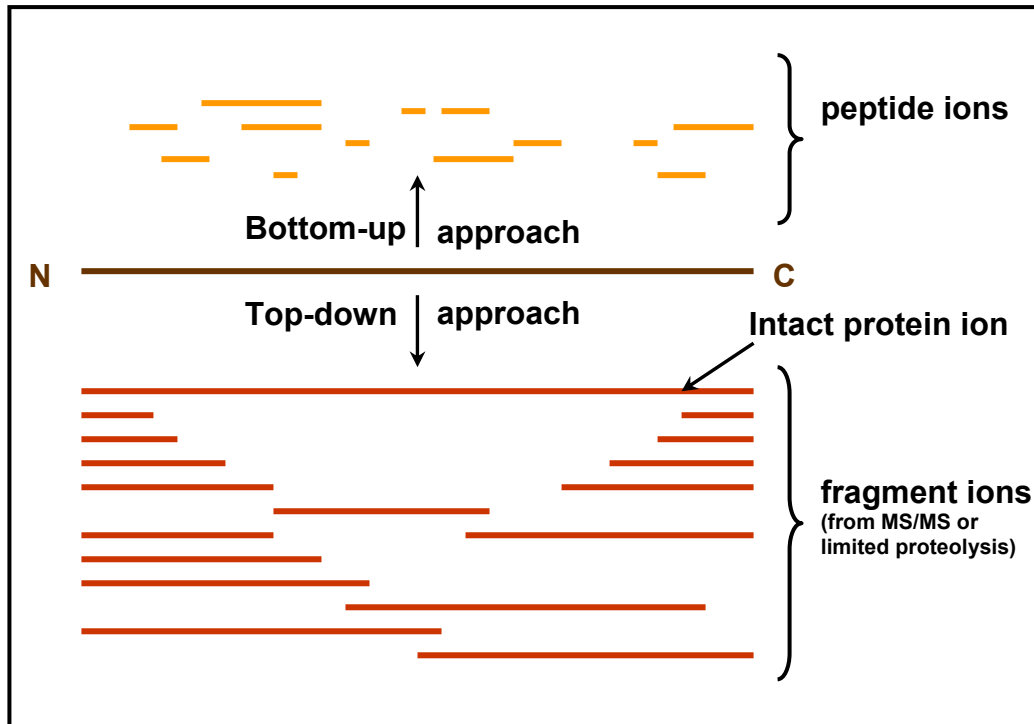


Figura 3.1 Confronto tra le strategie bottom-up e top-down per l'identificazione delle proteine. Figura adattata da [112].

3.1 Identificazione "top-down"

La metodica "top-down" prevede l'utilizzo di una proteina intatta che viene frammentata direttamente all'interno dello spettrometro di massa piuttosto che preventivamente digerita in soluzione. La molecola proteica viene tagliata così da creare una serie di ioni indicativi della sequenza. L'ESI-FT-ICR-MS viene per lo

più utilizzata in questo tipo di tecnica per la elevata risoluzione che garantisce in combinazione con l'ECD. Seppure questo tipo di approccio è in forte crescita e sviluppo esistono delle grandi limitazioni correlate alla mancanza di opportuni algoritmi necessari per l'interpretazione dei dati che ne riducono l'utilizzo.

3.2 Identificazione “bottom-up”

L'approccio di tipo “bottom-up” inizia invece con l'analisi dei peptidi derivanti dalla digestione delle proteine. I peptidi vengono utilizzati sia direttamente per rintracciare le molecole proteiche all'interno di specifici database (peptide mass fingerprinting), o frammentate per ottenere sequenze uniche per la proteina.

3.2.1 Peptide Mass Fingerprinting (PMF)

La metodica più diffusa per l'identificazione di una proteina si basa sulla rivelazione dell'impronta digitale peptidica o come più comunemente viene indicato in inglese Peptide Mass Fingerprinting (PMF) [113, 114]. In questo tipo di procedura la proteina di interesse, il più delle volte purificata attraverso elettroforesi bidimensionale, viene tagliata enzimaticamente o chimicamente e un'aliquota della miscela peptidica ottenuta è analizzata in MS. Le masse dei peptidi derivanti dalla digestione in-gel vengono acquisite e successivamente confrontate con una lista generata al computer ottenuta dalla digestione *in silico* di un database proteico o di un database contenente sequenze derivanti dalla traduzione di sequenze geniche. Le proteine o i prodotti genici che vengono indicati, attraverso l'ausilio del software di interrogazione delle banche dati, come i più favorevoli candidati per un'identificazione positiva della proteina incognita sono attentamente valutati. Un'identificazione attendibile di una catena polipeptidica è dipendente da numerosi parametri [115, 116]

- a) accuratezza nella determinazione della massa dei frammenti
- b) numero delle masse sottomesse per l'interrogazione del database
- c) distribuzione delle masse
- d) numero delle masse che sono in accordo tra quelle sperimentali e quelle teoriche per la proteina del database
- e) dimensioni del database di sequenza
- f) numero delle modificazioni a carico della proteina considerate

3.2.2 Digestione in-gel

Quasi tutti i protocolli di digestione in gel che vengono attualmente utilizzati nei laboratori dove si compiono analisi di proteomica in maniera routinaria si basano sulla procedura descritta da Rosenfold e collaboratori [117] che hanno rianalizzato i differenti step descritti dalle metodiche già esistenti. Tutti i lunghi procedimenti di lavaggio, liofilizzazione ed equilibratura [118, 119] necessari per eliminare il colorante, il sodio dodecil-solfato ed altri contaminanti associati al gel che possono ostacolare il processo di digestione dei gel colorati con blue Coomassie, sono stati sostituiti con un lavaggio al 50% di acetonitrile, essiccamento a temperatura ambiente e reidratazione con il tampone di digestione. Il protocollo di colorazione è stato modificato riducendo le concentrazioni di acido acetico e metanolo nella soluzione per la colorazione ed eliminando completamente lo stesso acido acetico nella soluzione decolorante. I peptidi che si ottengono utilizzando tale metodo possono essere facilmente recuperati per eluizione passiva fornendo una resa pari a circa 65% del materiale iniziale. Il materiale che in genere rimane corrisponde per lo più a peptidi di grandi dimensioni, superiori a 5 kDa. La metodica è stata anche adattata ai gel colorati con il silver stain [120] permettendo di avere a disposizione tecniche più fini e sensibili per la rivelazione e l'identificazione delle proteine. A questo ha fatto seguito la rimozione del silver stain [121, 122] e l'ottimizzazione dei tamponi di digestione utilizzando dimetil-formammide (DMF) [123] o n-octil-glucoside [124].

3.2.3 Software di interrogazione delle banche dati

Esistono un numero elevato di programmi disponibili in Internet che permettono l'interrogazione dei database. La complessità degli algoritmi che i programmi utilizzano varia da molto semplice, basandosi sulla comparazione diretta dei dati sperimentali con quelli teorici, ad elevata, quando si introducono variabili probabilistiche nel calcolo dei punteggi (score). I programmi più utilizzati sono: Mascot della Matrix-Science (<http://www.matrixscience.com>) e ProFound della Proteometrics (<http://65.219.84.5/service/prowl/profound/control.html>). Mascot funziona utilizzando l'algoritmo MOWSE, che tiene conto dell'abbondanza relativa dei peptidi nel database e delle dimensioni della proteina identificata. ProFound è basato sulla teoria Bayesiana [125]. Per poter utilizzare il sempre più crescente numero di sequenze genomiche a disposizione, i programmi di ricerca

sono stati adattati per eseguire le ricerche nelle banche dati di DNA utilizzando direttamente i dati ottenuti da esperimenti di PMF.

3.3 Sequenziamento peptidico

Se l'approccio tramite PMF per identificare una proteina fallisce, generalmente quello che si può tentare di fare è di sequenziare la catena polipeptidica attraverso esperimenti di massa tandem. Anche in questo caso esistono due differenti metodi che utilizzano i dati MS/MS ottenuti dai peptidi per il riconoscimento proteico. Il primo si avvale dell'uso dei dati MS/MS ottenuti sperimentalmente e non interpretati per l'interrogazione del database attraverso un confronto con i dati prodotti *in silico* dai peptidi proteolitici teorici [126, 127]. Nel secondo si interpretano invece manualmente i dati degli esperimenti di tandem MS realizzando un sequenziamento *de novo*.

3.3.1 Sequenziamento *in silico* (identificazione probabilità mediata)

La metodica si basa sull'uso di algoritmi che utilizzano sia dati non interpretati sia informazioni parziali di sequenza. Nel primo caso il software crea una lista di peptidi isobari al peptide sul quale la sequenza è stata effettuata. Per ognuno di questi peptidi candidati, sono calcolate le masse dei frammenti ionici ed è generato uno spettro MS/MS teorico. Lo spettro di massa tandem osservato è confrontato con quelli teorici e ad ogni confronto viene associato un punteggio che viene poi relazionato a tutti gli altri ottenuti. La limitazione maggiore di questo metodo è che l'algoritmo che determina l'assegnazione dello score fornisce sempre dei risultati che alcune volte possono rappresentare dei falsi positivi. Occorre quindi considerare che la qualità degli spettri deve essere molto elevata affinché l'algoritmo possa produrre un risultato attendibile. Un modo per poter ovviare a tale inconveniente consiste nell'interpretare manualmente i dati in maniera parziale così da fornire al software quelli che vengono indicati come marcatori di sequenza (sequence tag).

I marcatori di sequenza peptidica e il valore della loro massa, insieme con il valore della massa dello ione parente e con l'informazione sull'identità dell'amminoacido all'estremità C-terminale (legata all'uso di una proteasi specifica), permettono di eseguire una ricerca nei database garantendo unicamente l'identificazione di quel peptide specifico. Molto spesso per questo tipo

di indagine delle brevi sequenze di due o tre amminoacidi sono sufficienti per un'identificazione inequivocabile. Esistono però pur sempre delle limitazioni connesse alla difficoltà di poter automatizzare la procedura o all'inevitabile fallimento quando la massa del peptide è differente da quella presente nella banca dati.

3.3.2 Sequenziamento *de novo*

L'interpretazione manuale degli spettri MS/MS per l'identificazione di una proteina, viene utilizzata quando il sequenziamento *in silico* fallisce perchè si viene a creare la situazione in cui i dati che si sono ottenuti dagli spettri di massa non possono essere correlati con nessuna sequenza in banca dati. In questi casi, i dati ottenuti possono essere dovuti alla presenza di una nuova proteina, all'esistenza di varianti mutate di proteine conosciute, a modificazioni post-traduzionali o chimiche a carico della molecola proteica. Sebbene il sequenziamento dei peptidi *de novo* viene generalmente compiuto manualmente, diversi tentativi sono stati fatti per cercare di dedurre la sequenza peptidica direttamente dallo spettro ma il loro successo è stato ampiamente limitato perchè non è possibile differenziare i frammenti ionici che contengono l'originale regione N-terminale (ioni a, b, e c) da quelli contenenti la regione C-terminale (ioni x, y e z) [128]. E' per questo che alcuni approcci differenti atti a ridurre la complessità degli spettri MS/MS sono stati studiati.

Una delle procedure elaborate per ovviare alla complessità degli spettri di frammentazione peptidica prevede la parziale marcatura, attraverso isotopi stabili, del gruppo carbossilico al C-terminale utilizzando l' ^{18}O [129]. Tale metodica prevede che la digestione proteica venga realizzata in un tampone costituito da una miscela 1:1 di acqua contenente ^{18}O ed ^{16}O . Da questa marcatura risulta che nello spettro appaiono due picchi corrispondenti allo ione y per ogni frammento contenente la porzione C-terminale. Un picco rappresenta la naturale abbondanza di distribuzione dell'ossigeno, mentre il secondo mostra la distribuzione isotopica dell'ossigeno ^{18}O -marcato spaziato di 2 Da. La rilevanza di questo tipo di metodologia è limitata *in primis* dall'elevato costo del reagente isotopico ed anche dall'alta risoluzione delle masse, richiesta allo strumento che si utilizza, che deve essere in grado di distinguere uno shift di 2 Da del prodotto marcato.

Altre procedure alternative sono state studiate, molte delle quali prevedono la modificazione chimica dell'estremità N-terminale o C-terminale del peptide

attraverso l'aggiunta di una carica prestabilita. La derivatizzazione all'N-terminale con dei gruppi carichi positivamente porta ad un aumento nella rivelazione degli ioni a e b, mentre l'utilizzo di gruppi carichi negativamente che interagiscono con l'estremità C-terminale determina la produzione di sequenze costituite per lo più da ioni y e z. In generale possiamo dire che una grande varietà di composti sono stati testati per la derivatizzazione risultando più o meno soddisfacenti. Le maggiori limitazioni di gran parte di queste tecniche consistono in una bassa resa del processo, che si traduce nella necessità di avere a disposizione grandi quantità di peptide, lunghi tempi di reazione (2-16 h), largo eccesso del reagente e tamponi che determinano il bisogno di introdurre successivi step di clean-up del campione. In un panorama di questo tipo, due sono oggi le metodiche che destano maggiore interesse per la loro semplicità di applicazione, la riduzione dei tempi di reazione e l'accresciuta resa: la frammentazione chimicamente assistita (Chemically Assisted Fragmentation, CAF) e l'uso del 4-sulfofenil-isotiocianato (SPITC).

3.3.3 Frammentazione chimicamente assistita (CAF)

L'osservazione che i peptidi triptici con residui N-terminali fortemente acidi si frammentano con una grande resa in ioni y, ha spinto Keough e collaboratori [130-133] ad investigare sistematicamente nuove metodiche di modificazione. Dopo vari tentativi, si scoprì che la presenza all'N-terminale di un gruppo fortemente acido, come ad esempio un acido sulfonico, permetteva di aumentare la frammentazione del legame ammidico con la relativa formazione soprattutto di ioni b ed y. La carica negativa presente all'estremità N-terminale neutralizza la carica positiva normalmente localizzata a livello del C-terminale degli ioni b rendendo tali frammenti neutri e risultando in ultima analisi nella produzione di uno spettro di massa in cui sono esclusivamente presenti ioni y [130, 134]. Per poter avere una frammentazione di tipo ideale, utilizzando questa procedura assistita da sulfonazione, si richiede che i residui C-terminali siano arginina o lisina poichè tali amminoacidi sono basici e quindi carichi positivamente a pH neutro. Da questo ne consegue che tale metodo è ottimale per analizzare i frammenti ottenuti dalla digestione triptica che è di sicuro la metodica maggiormente applicata nelle analisi di proteomica su larga scala.

Un'altro importante step nell'ottimizzazione di questa procedura è stato fatto con la scoperta che la guanidinazione del ϵ -ammino gruppo dei peptidi terminanti in lisina ne determina la conversione in omoarginina. L'omoarginina è

decisamente molto meno basica della lisina e questo fa sì che la reazione di guanidinazione protegge ϵ -ammino gruppo da reazioni non desiderate con il reagente di sulfonazione. La reazione di guanidinazione aumenta il pattern di frammentazione dei peptidi terminanti in lisina. Il miglior agente per tale reazione di guanidinazione, prima dello step di sulfonazione, si è dimostrato essere l'O-metilisourea idrogeno solfato, poichè tale composto reagisce con ϵ -ammino gruppo dei residui laterali di lisina anche in presenza di ammine libere primarie all'N-terminale dei peptidi triptici. L'Amersham Biosciences ha commercializzato un kit per il sequenziamento CAF in cui il reagente di sulfonazione utilizzato, compatibile con i solventi acquosi, è costituito dall'acido 3-sulfopropionico NHS-estere. Il kit

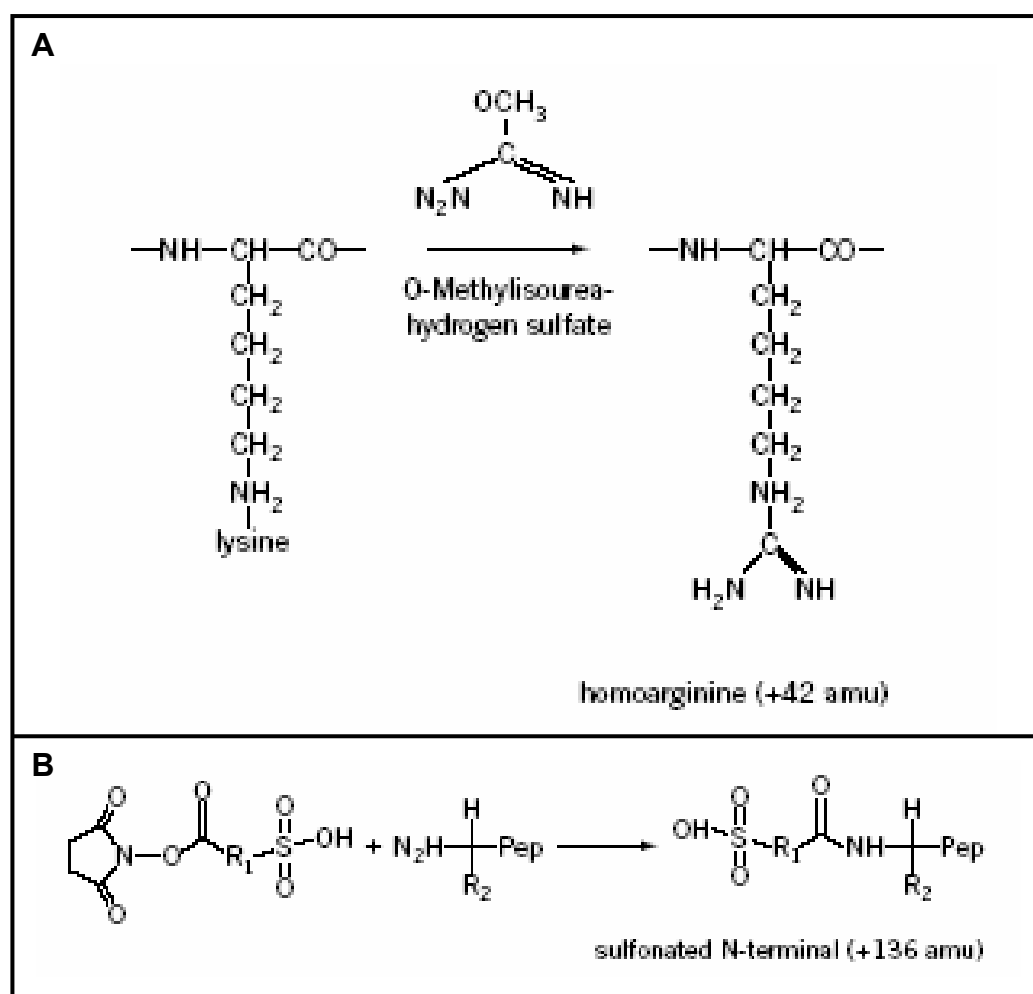


Figura 3.2 **A**, Guanidinazione dell' ϵ -amminogruppo della lisina. **B**, Sulfonazione dell'estremità N-terminale mediante il reagente CAF.

comprende anche il reagente di guanidinazione. Lo schema di reazione per la derivatizzazione CAF è mostrato in Fig. 3.2.

Alcuni amminoacidi possono dar origine a dei problemi quando si sequenziano dei peptidi derivatizzati con CAF. La presenza di prolina nella sequenza, ad esempio, determina un aumento del segnale dello ione y risultante dal taglio sul versante N-terminale del residuo, mentre il segnale per lo ione y derivante dal taglio sul versante C-terminale risulta diminuito [131, 135]. In maniera simile anche la frammentazione in corrispondenza della glicina dà origine ad un segnale più basso. In entrambi i casi, si deve però considerare che il segnale recupera intensità al successivo amminoacido risultando in una differenza di massa che corrisponde ad un dipeptide. Un altro tipo di problema si può avere quando un residuo di arginina o lisina è presente internamente alla sequenza peptidica ("miscleavage" triptico), il che rende la frammentazione meno specifica nel formare solo ioni b ed y . Infine è da sottolineare come spesso si possono rintracciare nello spettro degli ioni $y-17$, corrispondenti alla perdita di ammoniaca, invece dei normali ioni y [135].

3.3.4 Derivatizzazione con 4-sulfofenil-isotiocianato (SPITC)

Per lungo tempo la determinazione della sequenza amminoacidica di una proteina è stata realizzata attraverso la degradazione di Edman, una procedura chimica a tre step in cui i residui amminoacidici vengono tagliati uno alla volta a partire dall'estremità N-terminale ed identificati come derivati feniltioindantoinici (Pth). Nel primo step di tale procedura, definito di accoppiamento, l'estremità aminotermine è derivatizzata utilizzando il reagente fenil-isotiocianato (PITC). Poiché il PITC si è dimostrato un reagente ottimale per il sequenziamento attraverso la procedura di Edman si ebbe ragione di credere che il suo acido sulfonico SPITC (Fig. 3.3) potesse essere altrettanto efficiente nel derivatizzare ed

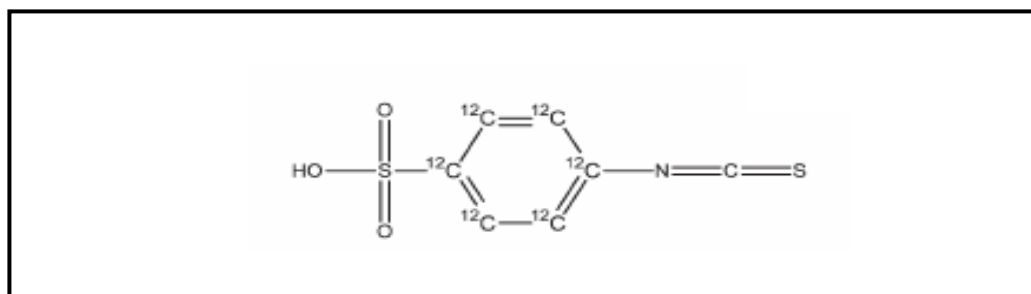


Figura 3.3 Struttura chimica del reagente di derivatizzazione SPITC.

assistere la frammentazione peptidica.

La veridicità di tale teoria fu dimostrata dal lavoro di Gevaert e collaboratori [136] che legarono covalentemente il gruppo carico negativamente all'estremità N-terminale dei peptidi. L'acquisizione degli spettri, realizzata in modalità positiva, ha permesso di rivelare soltanto i frammenti ionici carichi positivamente cioè quelli che si formano dalla ritenzione di carica all'estremità C-terminale. In accordo con quanto atteso, gli spettri MS/MS ottenuti presentavano per lo più ioni y ed alcuni dei loro ioni satellite z .

La reazione di derivatizzazione dell'N-terminale peptidico con lo SPITC è del tutto simile alla reazione di accoppiamento che utilizza il PITC nella degradazione di Edman. Considerando che il valore medio di pK_a dell'ammino gruppo terminale di un peptide è pari a 7.8 e che la reazione di accoppiamento necessita di un gruppo amminico, la procedura deve avvenire in ambiente alcalino (pH maggiore di 8). Nel tentativo di eliminare indesiderate reazioni secondarie di tipo competitivo occorre però evitare ambienti troppo basici (pH maggiore di 10). Tra le reazioni collaterali che si possono avere si può ricordare ad esempio l'idrolisi dello SPITC catalizzata da base, che oltre a determinare un consumo del reagente può dare origine a dei prodotti difficili da poter eliminare dal peptide accoppiato. Il gruppo ammino terminale ($\alpha-NH_2$) dei peptidi non è però l'unico che può reagire con lo SPITC. Anche i gruppi $\epsilon-NH_2$ dei residui di lisina interni alla sequenza peptidica reagiscono con lo SPITC. Poichè il pK_a del gruppo $\epsilon-NH_2$ dei residui di lisina è considerevolmente più alto del pK_a del gruppo $\alpha-NH_2$ del peptide, la reazione deve essere eseguita in un ambiente moderatamente basico (pH 8) così da indurre una reazione incompleta dei gruppi $\epsilon-NH_2$ che a tale valore di pH sono per lo più protonati. Affinchè la reazione dello SPITC con il gruppo $\alpha-NH_2$ sia completa sono richieste elevate concentrazioni del reagente e poichè la solubilità dello SPITC in acqua è bassa questo viene disciolto in miscele di solventi organici ed acqua. Non tutti i residui N-terminali dei peptidi reagiscono con la stessa velocità ed efficienza, ma la reazione può essere essenzialmente considerata giunta a completezza dopo trenta minuti a 50 °C utilizzando una soluzione di SPITC 2.5% (v/v) in solvente organico/acquoso. Tempi di reazione prolungati o temperature più elevate possono portare ad un aumento delle reazioni laterali. Prima del processo di derivatizzazione, come già visto per l'uso del CAF, può essere aggiunto uno step di guanidinazione per convertire i residui di lisina in omoarginina. La guanidinazione determina un aumento di 42 Da della massa del

peptide contenente lisina mentre l'aggiunta dello SPITC porta ad uno shift di massa di 214.97 Da.

CAPITOLO 4

BIBLIOGRAFIA

- [1] Wilkins, M. R., Sanchez, J. C., Gooley, A. A., Appel R. D., Umphery-Smith, I., Hochstrasser, D. F., Williams, K. L. (1995) Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotech. Gen. Eng. Rev.* **13**: 19-50.
- [2] Patterson, S. D., Aebersold, R. H. (2003) Proteomics: the first decade and beyond. *Nature* **33**: 311-323.
- [3] Kenyon, G. L., Demarini, D. M., Fuchs, E., Galas, D. J., Kirsch, J. F., Leyh, T. S., Moos, W. H., Petsko, G. A., Ringe, D., Rubin, G. M., Sheahan L. C. (2002) Defining the mandate of proteomics in the post-genomics era: workshop report. *Mol. Cell. Proteomics* **1**: 763-780.
- [4] Wilkins, M. R., Pasquali, C., Appel, R. D., Ou, K., Golaz, O., Sanchez, J. C., Jan, J. X., Gooley, A. A., Hughes, G., Umphery-Smith, L., Williams, K., L., Hochstrasser, D. F. (1996) From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Biotechnology* **14**: 61-65.
- [5] Celis, J., Ostergaard, M., Madsen, P., Honore, B., Dejgaard, K., Olsen, E., Vorum, H., Kristensen, D. B., Gromoval, I., Haunso, A., Van Damme, J., Puype, M., Vandekerckhove, J., Rasmussen, H. H. (1996) Human 2-D PAGE databases for proteome analysis in health and disease: <http://biobase.dk/cgi-bin/celis>. *FEBS Lett.* **398**: 129-134.
- [6] Anderson, N. G., Anderson, N. L. (1996) Twenty years of two-dimensional electrophoresis: past, present and future. *Electrophoresis* **17**: 443-453.
- [7] Pandey, A., Mann, M. (2000) Proteomics to study genes and genomes. *Nature* **405**: 837-846.
- [8] Aebersold, R., Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* **422**: 198-207.

- [9] Appel, R. D., Palagi, P. M., Walther, D., Vargas, J. R., Sanchez, J. C., Ravier, F., Pasquali, C., Hochstrasser, D. F. (1997) Melanie II--a third-generation software package for analysis of two-dimensional electrophoresis images: I. Features and user interface. *Electrophoresis* **18**: 2724-2734.
- [10] Corthals, G. L., Wasinger, V. C., Hochstrasser, D. F., Sanchez, J. C. (2000) The dynamic range of protein expression: a challenge for proteomic research. *Electrophoresis* **21**: 1104-1115.
- [11] Harry, J. L., Wilkins, M. R., Herbert, B. R., Packer, N. H., Aebersold, R. (2000) Proteomics: Capacity versus utility. *Electrophoresis* **21**: 1071-1081.
- [12] Gygi, S. P., Corthals, G. L., Zhang, Y., Rochon, Y., Aebersold, R. (2000) Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc. Natl. Acad. Sci. U.S.A.* **97**: 9390-9395.
- [13] Lose, J., Kobalz, U. (1995) Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome. *Electrophoresis* **16**: 1034-1059.
- [14] Hunt, D. F., Henderson, R. A., Shabanowitz, J., Sakaguchi, K., Michel, H., Sevilir, N., Cox, A. L., Appella, E., Engelhard, V. H. (1992) Characterization of peptides bound to the class I MCH molecule HLA-A2.1 by mass spectrometry. *Science* **255**: 1261-1263.
- [15] Yates III, J. R., McCormack, A. L., Schieltz, D., Carmack, E., Link, A. (1997) Direct analysis of protein mixtures by tandem mass spectrometry. *J. Prot. Chem.* **16**: 495-497.
- [16] Wolters, D. A., Washburn, M. P., Yates III, J. R. (2001) An automated multidimensional protein identification technology for shotgun proteomics. *Anal. Chem.* **73**: 5683-5690.
- [17] Washburn, M. P., Wolters, D., Yates III, J. R. (2001) Large scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **19**: 242-247.

- [18] Langen, H., Takacs, B., Evers, S., Berndt, P., Lahm, H. W., Wipf, B., Gray, C., Foutoulakis, M. (2000) Two-dimensional map of the proteome of *Haemophilus influenzae*. *Electrophoresis* **21**: 411-429.
- [19] Conrads, T. P., Anderson, G. A., Veenstra, T. D., Pasa-Tolic, L., Smith, R. D. (2000) Utility of accurate mass tags for proteome-wide protein identification. *Anal. Chem.* **72**: 3349-3354.
- [20] Smith, R. D., Pasa-Tolic, L., Lipton, M. S., Jensen, P. K., Anderson, G. A., Shen, Y., Conrads, T. P., Udseth, H. R., Harkewicz, R., Belov, M. E., Masselon, C., Veenstra, T. D. (2001) Rapid quantitative measurements of proteomes by Fourier transform ion cyclotron resonance mass spectrometry. *Electrophoresis* **22**: 1652-1668.
- [21] Smith, R. D., Anderson, G. A., Lipton, M. S., Pasa-Tolic, L., Shen, Y., Conrads, T. P., Veenstra, T. D., Udseth, H. R. (2002) An accurate mass tag strategy for quantitative and high throughput proteome measurements. *Proteomics* **2**: 513-523.
- [22] Jensen, P. K., Pasa-Tolic, L., Peden, K. K., Anderson, G. A., Tolic, N., Horner, J. A., Lipton, M. S., Bruce, J. E., Smith, R. D. (1999) Probing proteomes using capillary isoelectric focusing-electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **71**: 2076-2084.
- [23] Tonge, R., Shaw, J., Middleton, B., Rowlinson, R., Rayner, S., Young, J., Pognan, F., Hawkins, E., Currie, I., Davison, M. (2001) Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. *Proteomics* **1**: 377-396.
- [24] Unlu, M., Morgan, M. E., Minden, J. S. (1997) Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* **18**: 2071-2077.
- [25] Oda, Y., Huang, K., Cross, F. R., Cowburn, D., Chait, B. T. (1999) Accurate quantitation of protein expression and site specific phosphorylation. *Proc. Nat. Acad. Sci. U.S.A.* **8**: 6591-6596.

- [26] Chen, X., Smith, L. M., Bradbury, E. M. (2000) Site-specific mass tagging with stable isotopes in proteins for accurate and efficient protein identification. *Anal. Chem.* **72**: 1134-1143.
- [27] Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., Mann, M. (2002) Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as Simple and Accurate Approach to Expression Proteomics. *Mol. Cell. Proteomics* **1**: 376-386.
- [28] Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M., Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17**: 994-999.
- [29] Mirgorodskaya, O. A., Kozmin, Y. P., Titov, M. I., Korner, R., Sonksen, C. P., Roepstorff, P. (2000) Quantitation of peptides and proteins By Matrix-Assisted Laser Desorption/ionization Mass Spectrometry Using (18)-Labeled Internal Standards. *Rapid Commun. Mass. Spectrom.* **14**: 1226-1232.
- [30] Krishna, R. G., Wold, F. (1993) Post-translational modifications of proteins. *Adv. Enzymol. Relat. Areas Mol. Biol.* **67**: 265-298.
- [31] Mann, M., Jensen, O. N. (2003) Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* **21**(3): 255-261.
- [32] MacCoss, M. J., McDonald, W. H., Saraf, A., Sadygov, R., Clark, J. M., Tasto, J. J., Gould, K. L., Wolters, D., Washburn, M., Weiss, A., Clark, J. I., Yates III, J. R. (2002) Shotgun identification of protein modifications from protein complexes and lens tissue. *Proc. Natl. Acad. Sci. U.S.A.* **99**: 7900-7905.
- [33] Aebersold, R., Goodlett, D. R. (2001) Mass spectrometry in proteomics. *Chem. Rev.* **101**: 269-295.
- [34] Pandey, A., Podtelejnikov, A. V., Steen, H., Blagoev, B., Nielsen, M. M., Roche, S., Mann, M., Lodish, H. F. (2000) Analysis of receptor signalling pathways by mass spectrometry: identification of Vav-2 as substrate of epidermal and platelet-derived growth factor receptors. *Proc. Natl. Acad. Sci. U.S.A.* **97**: 179-184.
- [35] Andersson, L., Porath, J. (1986) Isolation of phosphoproteins by metal (Fe^{3+}) affinity chromatography. *Anal. Chem.* **154**: 250-254.

- [36] Stensballe, A., Andersen, S., Jensen, O. N. (2001) Characterization of phosphoproteins from electrophoretic gels by nano-scale FeIII affinity chromatography with off-line mass spectrometry analysis. *Proteomics* **1**: 207-222.
- [37] Ficcaro, S. B., McClelland, M. L., Stukenberg, P. T., Burke, D. J., Ross, M. M., Shabanowitz, J., Hunt, D. F., White, F. M. (2002) Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat. Biotechnol.* **20**: 301-305.
- [38] Larsen, M. R., Thingholm, T. E., Jensen, O. N., Roepstorff, P., Jorgensen, T. J. (2005) Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol. Cell. Proteomics* **4**: 873-886.
- [39] Grossert, J. S. (2001) A retrospective view of mass spectrometry and natural products – sixty years of progress, with a focus on contributions by R. G. Cooks. *Int. J. Mass Spectrom.* **212**: 65-79.
- [40] Dempster, A. J. (1921) Positive Ray Analysis of Lithium and Magnesium *Phys. Rev.* **18**: 415-422.
- [41] Neir, A. O. (1947) Electron impact mass spectrometry *Rev. Sci. Instrum.* **18**: 415-427.
- [42] Munson, M. S. B., Field, F. H. (1966) Chemical ionisation mass spectrometry. I. General introduction. *J. Am. Chem. Soc.* **88**: 2621-2630.
- [43] Beckey, H. D. (1969) Field desorption mass spectrometry: A technique for the study of thermally unstable substances of low volatility. *Int. J. Mass Spec. Ion Phys.* **2**: 500-503.
- [44] Benninghoven, A., Jaspers, D., Sichtermann, W. (1976) Secondary-ion emission of amino acids. *Appl. Phys.* **11**: 35-39.
- [45] Barber, M., Bordoli, R. S., Sedgwick, R. D., Tyler, A. N. (1981) Fast atom bombardment mass spectrometry of cobalamines. *Biomed. Mass Spectrom.* **8**: 492-495.
- [46] Sundqvist, B., Macfarlane, R. D. (1985) ²⁵²Cf-plasma desorption mass spectrometry. *Mass Spectrom. Rev.* **4**: 421-460.

- [47] Vestal, M. L. (2001) Methods of ion generation. *Chem. Rev.* **101**: 361-375.
- [48] Karas, M., Hillenkamp, F. (1988) Matrix-assisted laser desorption/ionisation, an experience. *Anal. Chem.* **60**: 2299-2301.
- [49] Tanaka, K., Waki, H., Ido, Y., Akita, S. (1988) Protein and polymer analyses up to m/z 100000 by laser ionization time-of-flight mass spectrometry. *Rapid Comm. Mass Spectrom.* **2**: 151-153.
- [50] Yamashita, M., Fenn, J. B. (1984) Electrospray ion source. Another variation on the free-jet theme. *J. Phys. Chem.* **88**: 4451-4459.
- [51] Jonsson, G. G. P., Hedin, A. B., Håkansson, P. L., Sundqvist, B. U. R., Säve, B. G.-S., Nielsen, P. F., Roepstorff, P., Johansson, K.-E., Kamensky, I., Lindberg, S. L. (1986) Plasma desorption mass spectrometry of peptides and proteins absorbed on nitrocellulose. *Anal. Chem.* **58**: 1084-1087.
- [52] Alai, M., Demirev, P. Fenselau, C., Cotter, R. J. (1986) Glutathione as matrix for plasma desorption mass spectrometry of large peptides. *Anal. Chem.* **58**: 1303-1307.
- [53] Mumma, R. O., Vastola, F. J. (1972) Analysis of organic salts by laser ionization mass spectrometry-sulfonates, sulfates and thiosulfates. *Org. Mass. Spectrom.* **6**: 1373-1376
- [54] Hardin, E. D., Vestal, M. L. (1981) Laser ionisation mass spectrometry of non-volatile samples. *Anal. Chem.* **53**: 1492-1497.
- [55] Beavis, R. C., Chait, B. T. (1989) Matrix assisted laser desorption mass spectrometry using 355 nm radiation. *Rapid Commun. Mass Spectrom.* **3**: 436-438.
- [56] Karas, M., Bahr, U., Ingendoh, A., Nordhoff, E. (1990) Matrix-assisted laser-desorption mass spectrometry using 355 nm radiation. *Anal. Chem. Acta.* **241**: 175-185.
- [57] Beavis, R. C., Chait, B. T. (1989) Cinnamic acid derivatives as matrices for ultraviolet laser desorption mass spectrometry of proteins. *Rapid Commun. Mass Spectrom.* **3**: 432-435.
- [58] Strupat, K. K., Karas, M., Hillenkamp, F. (1991) 2,5-Dihydroxybenzoic acid: a

new matrix for laser desorption-ionization mass spectrometry. *Int. J. Mass. Spectrom. Ion Processes* **111**: 89-102.

[59] Beavis, R. C., Chaudhary, T., Chait, B. T. (1992) Alpha-cyano-4-hydroxycinnamic acid as a matrix for matrix assisted laser desorption mass spectrometry. *Org. Mass Spectrom.* **27**: 156-158.

[60] Xiang, F., Beavis, R. C. (1994) A method to increase contaminant tolerance in protein matrix-assisted laser desorption ionisation by the fabrication of thin protein-doped polycrystalline films. *Rapid. Commun. Mass Spectrom.* **8**: 199-204.

[61] Vorm, O., Roepstorff, P., Mann, M. (1994) Improved resolution and very high sensitivity in MALDI TOF of matrix surfaces made by fast evaporation. *Anal. Chem.* **66**: 3281-3287.

[62] Karas, M. Bachmann, D., Hillenkamp, F. (1985) Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules. *Anal. Chem.* **57**: 2935-2939.

[63] Karas, M., Bachmann, D., Bahr, U., Hillenkamp, F. (1987) Matrix-assisted ultraviolet-laser desorption of nonvolatile compounds. *Int. J. Mass Spectrom. Ion Processes* **78**: 53-68.

[64] Zenobi, R. Knochenmuss, R. (1998) Ion formation in MALDI mass spectrometry. *Mass Spectrom. Rev.* **17**: 337-366.

[65] Gluckmann, M., Karas, M. (1999) The initial ion velocity and its dependence on matrix, analyte and preparation method in ultraviolet matrix-assisted laser desorption/ionization. *J. Mass Spectrom.* **34**: 467-477.

[66] Rosinke, B., Strupat, K., Hillenkamp, F., Rosenbusch, J. J. (1995) Matrix-assisted laser desorption/ionization mass spectrometer (MALDI-MS) of membrane-proteins and noncovalent complexes. *Mass Spectrom.* **30**: 1462-1468.

[67] Pfenninger, A., Finke, B., Stahl, B., Sawatzki, G. J. (1999) Matrix optimization for matrix-assisted laser desorption/ionization mass spectrometry of oligosaccharides from human milk. *Mass Spectrom.* **34**: 98-104.

- [68] Preston, L. M., Murray, K. K., Russel, D. H. (1993) Reproducibility and quantitation of Matrix assisted Laser Desorption ionization mass spectrometry: effect of nitrocellulose on peptide ion yields. *Biol. Mass Spectrom.* **22**: 544-550.
- [69] Gusev., A. I., Wilkinson, W. R., Proctor, A., Hercules, D. M. (1995) Improvement of signal reproducibility and matrix/comatrix effects in MALDI analysis. *Anal. Chem.* **67**: 1034-1041.
- [70] Karas, M., Gluckmann, M., Schafer, J. J. (2000) Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors. *J. Mass Spectrom.* **35**: 1-12.
- [71] Hillenkamp, F., Karas, M., Beavis, R. C., Chait, B. T. (1991) Matrix-Assisted Laser Desorption/Ionization mass spectrometry of biopolymers. *Anal. Chem.* **63**: 1193-1203.
- [72] Dai, Y. Q., Whittall, R. M., Li, L. (1996) Confocal Fluorescence Microscopic Imaging for Investigation the Analyte Distribution in Maldi Matrices. *Anal. Chem.* **28**: 2494-2500.
- [73] Axelsson, J., Hoberg, A. M., Waterson, C., Myatt, P., Shield, G. L., Verney, J., Haddleton, D. M., Derrick, P. J. (1997) Improved reproducibility and increased signal intensity in matrix-assisted laser desorption/ionization as a result of electrospray sample preparation. *Rapid Commun. Mass Spectrom.* **11**: 209-213.
- [74] Mormann, M., Bashir, S., Derrick, P. J., Kuck, D. (2000) Gas-phase basicities of the isomeric dihydroxybenzoic acids and gas-phase acidities of their radical cations. *J. Am. Soc. Mass Spectrom.* **11**: 544-552.
- [75] Griffin, T. J., Gygi, S. P., Rist, B., Aebersold, R., Loboda, A., Jilkine, A., Ens, W., Standing, K. G. (2001) Quantitative proteomic analysis using a MALDI quadrupole time-of-flight mass spectrometer. *Anal. Chem.* **73**: 978-986.
- [76] Dole, M., Mack, L. L., Hines, R. L., Mobley, R. C., Ferguson, L. D., Alice, M. B. (1968) Molecular beams of macroions. *J. Chem. Phys.* **49**: 2240-2249.
- [77] Mann, M., Meng, C. K., Fenn, J. B. (1988) Parent mass information from sequences of peaks of multiply charged ions. *Proc. 36th ASMS Conf. Mass Spectrom.* Allied Topics, San Francisco, CA. 1207-1208.

- [78] Gaskell, S. J. (1997) Electrospray: principles and practice. *J. Mass Spectrom.* **32**: 677-688.
- [79] Iribane, J. V., Thomson, B. A. (1976) On the evaporation of small ions from charged droplets. *J. Chem. Phys.* **64**: 2287-2294.
- [80] Thomson, B. A., Iribane, J. V. (1979) Field induced ion evaporation from liquid surfaces at atmospheric pressure. *J. Chem. Phys.* **71**: 4451-4463.
- [81] Schmelzeisen-Redecker, G., Butfering, L., Rollgen, F. W. (1989) Desolvation of ions and molecules in thermospray mass spectrometry. *Int. J. Mass Spectrom. Ion. Proc.* **90** : 139-150.
- [82] Cole, R. B. (2000) Some tenets pertaining to electrospray ionization mass spectrometry. *J. Mass Spectrom.* **35**: 763-772.
- [83] Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., Whitehouse, C. M. (1989) Electrospray ionisation for mass spectrometry of large biomolecules. *Science* **246**: 64-71.
- [84] Gale, D. C., Smith, R. D. (1993) Small volume and flow-rate electrospray ionisation mass spectrometry of aqueous samples. *Rapid. Commun. Mass Spectrom.* **7**: 1017-1021.
- [85] Emmett, M. R., Caprioli, R. M. (1994) Micro-electrospray mass spectrometry: ultra-high sensitivity analysis of peptides and proteins. *J. Am. Soc. Mass Spectrom.* **5**: 605-613.
- [86] Wilm, M. S., Mann, M. (1994) Electrospray and Taylor-cone theory, Dole's beam of macromolecules at last? *Int. J. Mass Spectrom. Ion Proc.* **136**: 167-180.
- [87] Wilm, M. S., Mann, M. (1996) Analytical properties of the nanoelectrospray ion source. *Anal. Chem.* **68**: 1-8.
- [88] Geromanos, S., Freclتون, G., Tempst, P. (2000) Tuning of an electrospray ion source for maximum peptide-ion transmission into mass spectrometer. *Anal. Chem.* **72**: 777-790.

- [89] Stephens, W. E. (1946) A pulsed mass spectrometer with time dispersion. *Phys. Rev.* **69**: 691.
- [90] Cameron, A. E., Eggers, D. F. (1948) An ion "Velocitron" *Rev. Sci. Instrum.* **19**: 605-607.
- [91] Wiley, W. C., McLaren, I. H. (1955) Time-of-flight mass spectrometer with improved resolution. *Rev. Sci. Instrum.* **26**: 1150-1157.
- [92] Whittall, R. M., Li, L. (1995) High-Resolution matrix Assisted Laser Desorption Ionization In A Linear Time-of-Flight Mass-Spectrometer. *Anal. Chem.* **67**: 1950-1954.
- [93] Brown, S. R. Lennon, J. J. (1995) Mass resolution improvement by incorporation of pulsed ion extraction in a matrix assisted laser desorption/ionisation linear time-of-flight mass-spectrometer. *Anal. Chem.* **67**: 1998-2003.
- [94] Vestal, M. P., Juhasz, P., Martin, S. A. (1995) Delayed Extraction Matrix-Assisted Laser Desorption Time-of-Flight Mass-Spectrometry. *Rapid Commun. Mass. Spectrom.* **9**: 1044-1050.
- [95] Mamyrin, B. A., Karataev, V. I., Shmikk, D. V., Zagulin, V. A. (1973) The linear mass reflectron. *Sov. Phys. JETP.* **37**: 45.
- [96] Cotter, R. J. (1992) Time-of-flight mass spectrometry for the structural analysis of biological molecules. *Anal. Chem.* **64**: 1027A-1039A.
- [97] Cotter, R. J. (1999) The New Time-of-Flight Mass Spectrometry. *Anal. Chem.* **71**: 445A-451A.
- [98] Mamyrin, B. A. (2001) Time-of-Flight mass spectrometry (concepts, achievements and prospects). *Int. J. Mass Spectrom. Ion Proc.* **206**: 251-266.
- [99] Guilhaus, M., Selby, D., Mlynski, V. (2000) Orthogonal acceleration time-of-flight mass spectrometry. *Mass Spectrom. Rev.* **19**: 65-107.

- [100] Bienvenut, W. V., Deon, C., Pasquarello, C., Campbell, J. M., Sanchez, J. C., Vestal, M. L., Hochstrasser, D. F. (2002) Matrix-assisted laser desorption/ionization-tandem mass spectrometry with high resolution and sensitivity for identification and characterization of proteins. *Proteomics* **2**: 868-876.
- [101] Paul, W., Raether, M. (1955) Das elektrische Massenfiter. *Z. Phys.* **140**: 262-273.
- [102] March, R. E. (1997) An Introduction to Quadrupole Ion Trap Mass Spectrometry. *J. Mass Spectrom.* **32**: 351-369.
- [103] Jennings, K. R (1968) Collision-Induced Decompositions of Aromatic Molecular Ions. *Int. J. Mass Spec. Ion Phys.* **1**: 227-235.
- [104] Haddon, W. F., McLafferty, F. W. (1968) Metastable Ion Characteristics. Collision-Induced Metastables. *J. Am Chem. Soc.* **90**: 4745-4746.
- [105] Roepstorff, P., Fohlman, J. (1984) Proposal for A Common Nomenclature for Sequence Ions In Mass-Spectra of Peptides. *Biomed. Mass Spectrom.* **11**: 601-601.
- [106] Biemann, K. (1988) Contributions of Mass-Spectrometry to Peptide and Protein- Structure. *Biomed. Environ. Mass Spectrom.* **16**: 99-111.
- [107] Morris, H. R., Paxton, T., Dell, A., Langhorne, J., Berg, M., Bordoli, R. S., Hoyes, J., Bateman, R. H. (1996) High sensitivity collisionally-activated decomposition tandem mass spectrometry on a novel quadrupole/orthogonal-acceleration time-of-flight mass spectrometer. *Rapid Commun. Mass Spectrom.* **10**: 889-896.
- [108] March, R. E., Hughes, R. J. (1989) Quadrupole Storage Mass Spectrometry. In. Pp. 332-333, John Wiley and Sons, New York.,
- [109] Amster, I. J. (1996) Fourier Transform Mass Spectrometry. *J. Mass Spectrom.* **31**, 1325-1337.

- [110] Marshall, A. G., Hendrickson, C. L., Jackson, G. S. (1998) Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: A Primer. *Mass Spectrom. Rev.* **17**, 1-35.
- [111] Kettman, J. R., Frey, J. R., Lefkovits, I. (2001) The proteome, transcriptome and genome: top down or bottom up analysis? *Biomol. Eng.* **18**: 207-212.
- [112] Forbes, A. J., Mazur, M. T., Patel, H. M., Walsh, C. T., Kelleher, N. L. (2001) Toward efficient analysis of < 70 kDa proteins with 100% sequence coverage. *Proteomics* **1**: 927-933.
- [113] Henzel, W. J., Billeci, T. M., Wong, J. T., Grimley, S. C., Watanabe, C. (1993) Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence database. *Proc. Natl. Acad. Sci. U.S.A.* **90**: 5011-5015.
- [114] Henzel, W. J., Watanabe, C., Stults, J. T. (2003) Protein identification: the origins of peptide mass fingerprinting. *J. Am. Soc. Mass Spectrom.* **14**: 931-942.
- [115] Lahm, H. W., Langen, H. (2000) Mass spectrometry: a tool for identification of proteins separated by gels. *Electrophoresis* **21**: 2105-2114.
- [116] Jensen, O. N., Larsen, M. R., Roepstorff, P. (1998) Mass spectrometric identification and microcharacterization of proteins from electrophoretic gels: strategies and applications. Proteins: Structure, Function, and Genetics. *Proteins Suppl. 2* **999**: 74-89.
- [117] Rosenfeld, J., Capdeville, J., Guillemot, J. C., Ferrara, P. (1992) In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. *Anal. Biochem.* **203**: 173-179.
- [118] Eckerskorn, C., Lottspeich, F. (1989) Internal amino acid sequence analysis of proteins separated by gel electrophoresis after tryptic digestion in polyacrylamide matrix. *Chromatographia* **28**: 92-94.
- [119] Simpson, R. J., Moritz, R. L., Nice, E. C., Grego, B. (1987) A high-performance liquid chromatography procedure for recovering subnanomole amounts of protein from SDS-gel electroeluates for gas-phase sequence analysis. *Eur. J. Biochem.* **165**: 21-29.

- [120] Shevchenko, A., Wilm, M., Vorm, O., Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* **68**: 850-858.
- [121] Gharahdaghi, F., Weiberg, C. R., Meagher, D. A., Imai, B. S., Mische, S. M. (1999) Mass spectrometric identification of proteins from silver-stained polyacrylamide gel: A method for the removal of silver ions to enhance sensitivity. *Electrophoresis* **20**: 601-605.
- [122] Sumner, L. W., Wolf-Sumner, B., White, S. P., Asirvatham, V. S. (2002) Silver stain removal using H₂O₂ for enhanced peptide mass mapping by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **16**: 160-168.
- [123] Soskic, V., Godovac-Zimmermann, J. (2001) Improvement of an in-gel tryptic digestion method for matrix-assisted laser desorption/ionization time-of-flight mass spectrometry peptide mapping by use of volatile solubilizing agents. *Proteomics* **11**: 1364-1367.
- [124] Katayama, H., Nagasu, T., Oda, Y. (2001) Improvement of in-gel digestion protocol for peptide mass fingerprinting by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **15**: 1416-1421.
- [125] Zhang, W., Chait, B. T. (2000) ProFound: an expert system for protein identification using mass spectrometric peptide mapping information. *Anal. Chem.* **72**: 2482-2489.
- [126] Eng, J., McCormack, A., Yates III, J. (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **5**: 976-989.
- [127] Perkins, D. N., Pappin, D. J. C., Creasy, D. M., Cottrell, J. S. (1999) Probability-based protein identification by searching sequence databases using mass spectrometric data. *Electrophoresis* **20**: 3551-3567.

- [128] Taylor, J. A., Johnson, R. S. (2001) Implementation and uses of automated de novo peptide sequencing by tandem mass spectrometry. *Anal. Chem.* **73**: 2594-2604.
- [129] Shevchenko, A., Wilm, M., Mann, M. (1997) Peptide sequencing by mass spectrometry for homology searches and cloning of genes. *J. Prot. Chem.* **16**: 481-490.
- [130] Keough, T., Ypungquist, R. S., Lacey, M. P. (1999) A method for high sensitivity peptide sequencing using postsource decay matrix-assisted laser desorption ionization mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **96**: 7131-7136.
- [131] Keough, T., Lacey, M. P., Youngquist, R. S. (2000) Derivatization procedures to facilitate *de novo* sequencing of lysine-terminated tryptic peptides using postsource decay matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **14**: 2348-2356.
- [132] Keough, T., Lacey, M. P., Youngquist, R. S. (2002) Solid-phase derivatization of tryptic peptides for rapid protein identification by matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **16**: 1003-1015.
- [133] Keough, T., Youngquist, R. S., Lacey, M. P. (2003) Sulfonic Acid Derivatives for Peptide Sequencing by MALDI MS. *Anal. Chem.* **75**: 156A-165A.
- [134] Wang, D., Kalb, S. R., Cotter, R. J. (2004) Improved procedures for N-terminal sulfonation of peptides for matrix-assisted laser desorption/ionization post-source decay peptide sequencing. *Rapid Commun. Mass Spectrom.* **18**: 96-102.
- [135] Hellman, U., Bhikhabhai, R. (2002) Easy amino acid sequencing of sulfonated peptides using post-source-decay on matrix-assisted laser desorption/ionization time-of-flight mass spectrometer equipped with a variable voltage reflector. *Rapid Commun. Mass Spectrom.* **16**: 1851-1859.

[136] Chen, P., Nie, S., Mi, W., Wang, X., Liang, S. (2004) *De novo* sequencing of tryptic peptides sulfonated by 4-sulfophenyl-isothiocyanate for unambiguous protein identification using post-source decay matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **18**: 191-198.

PARTE II

**“Utilizzo di una strategia di indagine di
tipo proteomico per lo studio
dell'apparato fotosintetico nelle piante
superiori”**

REGULAR ARTICLE

A proteomic approach for investigation of photosynthetic apparatus in plants

Corrado Ciambella¹, Peter Roepstorff², Eva Mari Aro³ and Lello Zolla¹

¹ Department of Environmental Sciences, University of Tuscia, Viterbo, Italy

² Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense, Denmark

³ Plant Physiology and Molecular Biology, Department of Biology, University of Turku, Turku, Finland

The proteome of the photosynthetic apparatus of barley (*Hordeum vulgare*), obtained by analysis of thylakoids without any previous fractionation, was mapped by native electrophoresis followed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) as the second dimension two-dimensional-blue native (2-D/BN)/SDS-PAGE). This protocol provided an excellent alternative to the 2-D-isoelectric focusing/sodium dodecyl sulfate-polyacrylamide gel electrophoresis for 2-D separation of the most hydrophobic thylakoid proteins. Monocots and dicots showed significant differences in the first dimension while in the second dimension patterns appeared similar. Identification of each spot was performed by internal peptide primary sequence determination using both nano-electrospray ionization tandem mass spectrometry and, to a lesser extent, peptide mass fingerprinting matrix-assisted laser desorption/ionization-time of flight using MALDI-TOF. This is due in particular to the fact that a limited number of peptides was obtained after trypsin digestion of these highly hydrophobic proteins. A larger number of peptides from hydrophilic intermembrane domains of transmembrane proteins were detected. Despite this, about 70% of the expected proteins were identified, including proteins with grand average of hydropathicity scores higher than 0.5. It is therefore reasonable to assert that protein hydrophobicity is not the limiting factor. Small proteins were not well identified with trypsin digestion. Instead some of these could be identified using acid hydrolysis. The method presented here does not require prefractionation of different thylakoid complexes and consequently gives confidence in comparing the proteome of the photosynthetic apparatus before and after treatment. It thus allows us to understand the molecular mechanisms underlying physiological adaptations of higher plants and to perform screening of photosynthetic mutants.

Received: July 9, 2004
Revised: October 12, 2004
Accepted: November 11, 2004

**Keywords:**

Blue native gel / Limited acid hydrolysis / Peptide mass fingerprinting / Tandem mass spectrometry / Thylakoid

Correspondence: Prof. Dr. Lello Zolla, University of Tuscia, Largo dell'Università, 01100 Viterbo, Italy

E-mail: zolla@unitus.it

Fax: +39-0761-357-630

Abbreviations: BNG, blue native gel; Cyt. b_6/f , cytochrome b_6/f complex; DM, n-dodecyl- β -D-maltoside; FA, formic acid; GRAVY, grand average of hydropathicity; LHC, light-harvesting complex; PS, photosystem

1 Introduction

Oxygenic photosynthesis takes place in thylakoid membranes in cyanobacteria, green algae and higher plants, where multiprotein complexes convert the energy from sunlight into chemical energy. The four protein complexes photosystem I (PSI), photosystem II (PSII), ATP-synthase and cytochrome b_6/f complex (Cyt. b_6/f) consist of at least 70 different proteins [1]. A great number of these proteins are

highly hydrophobic membrane integral-helical protein subunits which are embedded in the membrane system. PSII contains a core complex formed by a heterodimeric structure involving PsbA and PsbD associated with at least 20 larger and smaller proteins [2]. This is surrounded by a specific light-harvesting system, which is known as the major protein-chlorophyll *a/b* complex (LHCII), as well as minor antennae called Lhcb4, Lhcb5 and Lhcb6 [3]. PSI also contains a central heterodimer of two subunits called PsaA and PsaB which interact with 11 membrane proteins. This core complex is surrounded by four different polypeptides of the light-harvesting complex (LHCI) [4]. ATP-synthase is composed of two main portions: a hydrophilic one (F_1) which is an oligomerization of five different proteins and a membrane-integrated part (CF_0) consisting of four different subunits [5]. Finally, there is the Cyt. b_6/f complex which contains nine hydrophobic proteins [6].

Membrane proteomes are especially challenging and thus very few membrane proteomes have been analyzed successfully. Thus analysis of these large and highly hydrophobic protein complexes is of interest. However, despite the considerable arsenal of available technologies, analytical protocols for membrane proteins are still divergent and require a substantial effort of adaptation and optimization to investigate particular problems. Moreover, in both photosystems, the amino acid composition of the antenna proteins is highly similar in both length and sequence [3], and therefore the characterization of these proteins requires tedious electrophoresis and immunoblotting procedures [7, 8]. Recently, separation of thylakoid proteins has been attempted by reversed phase HPLC coupled with MS for determination of intact M_r via MS analysis [9–12] avoiding loss of protein by PAGE and enzymatic digest. The applicability of this method has been demonstrated lately, allowing detection of approximately 90 intact mass tags corresponding to 40 gene products with variable PTMs [13]. However, because of the high complexity of the photosynthetic apparatus, a complete and simultaneous separation and identification of all or most thylakoid proteins cannot be performed by chromatographic techniques without a preliminary preseparation by sucrose gradient ultracentrifugation.

2-DE [14] is most suitable for separating and displaying a great number of proteins. With this technique, large protein complexes, showing a wide size range (10 kDa to 500 kDa), and both moderately hydrophobic or very acidic/basic proteins can be isolated and visualized. IEF (first dimension) followed by SDS-PAGE (second dimension) provided an effective way to resolve large numbers of soluble and peripheral membrane proteins [15, 16], but is not useful for isolation of native protein complexes or separation of highly hydrophobic membrane proteins [17, 18]. These proteins would hardly stay in solution upon approaching their *pI* value and would have rather similar *pIs*, due to the limited number of amino acids with charged side chains. In a previous study Hippler *et al.* [19] presented a 2-D protein map from *Chlamydomonas reinhardtii* (green algae) which also

included some hydrophobic thylakoid membrane proteins (e.g. PsaA (PSI) and PsbA (PSII)). Recently an extensive analysis of the *Arabidopsis thaliana* peripheral and integral membrane proteome was performed by sequential extractions with organic solvents followed by multidimensional protein separation steps [1]. This study resulted in identification of a large number of proteins and it also revealed that the 2-D gel system using IEF in the first dimension has only limited capacity compared to preextraction of proteins by organic solvents and subsequent monodimensional SDS-PAGE.

An attractive technology for first dimension gel separations is blue native gel (BNG) [20] where solubilized thylakoid membranes can be loaded with a dye in order to charge the complexes for separation by PAGE. In the second dimension protein complexes may then be resolved by SDS-PAGE into single protein subunits [21–23]. This system, originally applied to mitochondrial inner membranes [24], has been recently applied both to *Synechocystis* thylakoids allowing functional proteomic analysis [25] and to *Chlamydomonas*, even though a low number of proteins were identified in both cases due to the use of only the MALDI approach [26]. A recent paper from Heinemeyer *et al.* [27] also used this approach to investigate the supramolecular organization of photosystems in *Arabidopsis*.

In this paper 2-D BNG/SDS-PAGE has been applied to integral thylakoid proteins from higher plants. Gel spots were detected using ESI-MS/MS and MALDI-TOF. About 70% of the total proteins expected were identified. Unfortunately, in the case of highly hydrophobic transmembrane proteins, especially short-chain ones, identification by MS is hampered by the fact that most proteins when digested by trypsin produce a limited number of peptides useful for analysis and by the lack of protein sequences of barley in databases. On the contrary most proteins were successfully identified by peptide sequencing of hydrophilic intermembrane peptides using ESI-MS/MS and in some cases we overcame the limits of tryptic in-gel digestion using alternative tools for proteins fragmentation (e.g. limited acid hydrolysis).

2 Materials and methods

2.1 Extraction of thylakoid membranes

Chloroplast thylakoid membranes were isolated from leaves of barley (*Hordeum vulgare*) and tomato (*Lycopersicon esculentum*) according to the method of Berthold *et al.* [28] with modifications reported elsewhere [29]. Leaves were collected at night in the dark and at different periods of the year to minimize any seasonal and light effects. To allow comparison of the results, the experimental conditions for extraction were kept constant for both species examined. All steps were performed in dim light at 4°C.

2.2 BNG-PAGE/SDS-PAGE

BNG-PAGE of integral thylakoid proteins was performed according to Kügler *et al.* [30], with the modifications of Suorsa *et al.* [31]. Thylakoid membranes were washed with washing buffer (330 mM sorbitol, 50 mM BisTris-HCl, pH 7.0, and 250 µg/mL Pefabloc as a protease inhibitor (Roche, Indianapolis, IN, USA), collected by centrifugation ($3500 \times g$ for 2 min at 4°C), and resuspended in 25BTH20G (20% w/v glycerol, 25 mM BisTris-HCl, pH 7.0, and 250 µg/mL Pefabloc) at a chlorophyll concentration of 1 mg/mL. An equal volume of resuspension buffer containing 2% or 4% w/v *n*-dodecyl β -D-maltoside (DM; Sigma, St. Louis, MO, USA) was added under continuous mixing, and the solubilization of membrane protein complexes was allowed to occur for 3 min on ice. Insoluble material was removed by centrifugation at $18\,000 \times g$ for 15 min. The supernatant was mixed with 0.1 volumes of Coomassie Blue solution (5% w/v Serva blue G, 100 mM BisTris-HCl, pH 7.0, 30% w/v sucrose, and 500 mM ϵ -amino-*n*-caproic acid) and loaded onto 0.75 mm thick 5% to 12.5% w/v acrylamide gradient gel (Hoefer Mighty Small mini-vertical unit; Amersham Biosciences, Uppsala, Sweden). Electrophoresis was performed at 4°C by increasing voltage gradually from 75 up to 200 V during the 4.5 h run. Moreover, we used a Protean II Bio-Rad (Hercules, CA, USA) electrophoresis system (180 $\times$ 160 mm, 0.75 mm thick) for the first dimension applying a constant voltage of 90 V overnight and gradually increasing this value up to 200 V the next day until the run was complete. For separation of proteins in the second dimension, the lanes of the BNG gel were excised and incubated in SDS sample buffer containing 5% v/v β -mercaptoethanol and 6 M urea for 30 min at room temperature. The lanes were then layered onto 1 mm thick SDS-PAGE gels [32], with 15% w/v acrylamide and 6 M urea in the separating gel. Proteins were separated overnight at 10 mA constant current *per* gel and 13°C. The proteins were visualized using Coomassie R-250, colloidal Coomassie and also by silver staining [33].

2.3 Determination of the chlorophyll concentration of isolated membranes

The chlorophyll content of the isolated membranes was determined according to the method of Porra [34].

2.4 In-gel digestion

Spots were carefully cut out of 2-D Silver-stained gels using a razor blade and subjected to in-gel trypsin digestion according to the method of Shevchenko *et al.* [35] with minor modifications. The gel pieces were swollen in a digestion buffer containing 50 mM NH_4HCO_3 and 12.5 ng/ μL trypsin (modified porcine trypsin, sequencing grade; Promega, Madison, WI, USA) in an ice bath. After 30 min the supernatant was removed and discarded, 20 μL of 50 mM NH_4HCO_3 was added and digestion was allowed to proceed at 37°C overnight. The supernatant containing tryptic peptides was dried

by vacuum centrifugation. Prior to mass spectrometric analysis, the peptide mixtures were redissolved in 10 μL of 5% formic acid (FA). The proteins in some selected spots were cleaved using limited acid hydrolysis [36]. The gel pieces were incubated for 1 h in 100 mM $\text{NH}_4\text{HCO}_3/\text{ACN}$ (1:1, v/v) to remove the Coomassie stain and then washed with water for another 30 min to remove ammonium bicarbonate. The water was discarded, and the gel pieces were shrunk with 50 μL of ACN. The ACN was discarded, and the gel pieces were dried in a vacuum centrifuge. Approximately 30 μL of 6 M HCl containing 10 mM DTT was added, and the sample was incubated for 45 min at 56°C. The sample was chilled to room temperature, and the gel pieces were spun in a bench-top centrifuge. A total of 25 μL of water was then added. After shaking briefly, the supernatant was recovered in a separate Eppendorf test tube (Eppendorf, Hamburg, Germany) and immediately dried in a vacuum centrifuge.

2.5 Protein Identification

2.5.1 PMF by MALDI-TOF-MS

Peptides were desalted and concentrated according to the method of Gobom *et al.* [37]. Homemade 5 mm nanocolumns were packed with chromatographic materials of increasing hydrophobicity in a constricted GELoader tip (Eppendorf). POROS R1, R2, oligo R3 resins (PerSeptive Biosystems, Framingham, MA, USA) and graphite powder were used in order to increase sequence coverage [38]. A syringe was used to force liquid through the columns by applying gentle air pressure. The columns were equilibrated with 20 mL of 5% FA and then analyte solutions were added. Peptide mixtures flushed through the first column were collected in the next one. This was repeated in the following order: R1, R2, oligo R3 and graphite powder. The columns were washed with 20 μL of 5% FA and the bound peptides subsequently eluted directly onto the MALDI target with 0.5 μL 2,5-dihydroxybenzoic acid solution (20 µg/ μL in ACN, 0.1% TFA, 70:30 vol/vol) or CHCA solution (20 µg/ μL in ACN, 0.1% TFA, 70:30 vol/vol). MALDI-TOF-MS was performed using a Voyager STR (PerSeptive Biosystems) equipped with delayed extraction. Positively charged ions were analyzed in reflectron mode, using delayed extraction. Typically, 100 shots were averaged to improve the S/N. MoverZ software (www.proteometrics.com) was used to analyze spectra. Spectra were calibrated using trypsin autolysis products (m/z 842.51 and 2211.10) as internal standards. In a few cases this was not possible and external calibration was carried out using peaks from a tryptic digest of β -lactoglobulin. Protein identification was performed by searching in a nonredundant protein sequence database (NCBIInr) using the MASCOT program (<http://www.matrixscience.com>). The following parameters were used for database searches: monoisotopic mass accuracy < 50 ppm, missed cleavages 1, complete carbamidomethylation of cysteines and partial oxidation of methionines. For positive identification, the score of the result of $[-10 \times \text{Log}(P)]$ had to be over the significance threshold level ($p < 0.05$).

2.5.2 Peptide sequencing by RP-HPLC-ESI-MS/MS

Peptide mixtures were separated using an Agilent system autosampler/nanoHPLC (Agilent technologies, Palo Alto, CA, USA). A sample volume of 8 μ L was loaded by the autosampler and concentrated on a commercial trap column (Zorbax 300 SB, C18, 5 μ m, 300 μ m id $\times$ 5 mm; Agilent Technologies) at a flow rate of 10 μ L/min. Separation was performed using a homemade capillary column (Zorbax 300 SB, C18, 3.5 μ m, 75 μ m id $\times$ 10 cm; Agilent Technologies) and a flow rate of 200 nL/min. HPLC solvents contained 0.1% FA and either 2% ACN (solvent A) or 98% ACN (solvent B). The column was pre-equilibrated with 100% solvent A. Elution was performed with a multisegment gradient with a first step from 0 to 5% solvent B in 5 min, a second one from 5 to 12% solvent B in 2 min, a third one from 12% to 50% solvent B in 30 min and a last segment from 50 to 100% solvent B in 8 min. Peptides were eluted directly onto a Bruker ion trap Esquire 3000+ (Bruker, Bremen, Germany). Capillary voltage was 1.5–2 kV and a dry gas flow rate of 3 L/min was used with a temperature of 230°C. The scan range used was from 300 to 1800 m/z . For protein identification the following parameters were used: complete carbamidomethylation of cysteines and partial oxidation of methionines, peptide mass tolerance $\pm$ 1.2 Da, fragment mass tolerance $\pm$ 0.9 Da, missed cleavages 2. In the case of proteins digested with limited acid hydrolysis desamidation of asparagine and glutamine to aspartic and glutamic acid was taken into account. For further investigations of some spots that didn't give direct identifications, off-line nano-ESI MS/MS was performed using a quadrupole-TOF instrument (Micromass, Manchester, UK) calibrated with NaI. When no proteins matched to given sequences the peptides were used for BLAST searches. To introduce short sequences to a blast search we used the fasta3 tool from the European Bioinformatics Institute (www.ebi.ac.uk/fasta33). Protein similarity searches were done at NCBI (<http://www.ncbi.nlm.nih.gov/blast/>).

3 Results

For each species examined, before extraction, leaves were harvested at night in the dark and frozen in liquid nitrogen to minimize any light effect and eliminate PTMs. Experimental conditions, such as incubation time and detergent concentrations for thylakoid protein extractions, were kept the same for all species whenever possible to ensure comparability of results.

3.1 First dimension: BNG-PAGE

A native PAGE system [20], was optimized for separation of thylakoid membranes. For this purpose sample solutions of thylakoid membranes were solubilized with DM, then mixed with Coomassie G-250 and loaded onto a PAGE system (first dimension) in order to separate intact protein complexes.

Figure 1 compares blue native separation of barley and tomato thylakoid membrane protein complexes, as representative species of monocots and dicots. Since the apparent molecular masses of the protein complexes in the 1-D native PAGE corresponded to the predicted molecular masses of all compounds, each complex may be recognized by its molecular mass. It can be seen that in tomato, as well as in all other dicots examined (data not shown), at least two supercomplexes containing PSI- and PSII-proteins with molecular masses over 600 kDa are present. In contrast, in barley and other monocots, the presence of supercomplexes can be revealed only upon Silver-staining, as shown in Fig. 1, suggesting a reduced organization of the photosynthetic apparatus or less stable supercomplexes in monocots. At about 550–600 kDa an intense green band appeared, which was assigned to PSI-LHCI complex and to a dimeric form of the PSII core complex. Both complexes comigrated due to their similar molecular mass. At 440 kDa a green band representing PSI/LHCI-intermediate appeared only in barley. At about 280 kDa, three Coomassie stained bands were closely adjoined and represent ATP-synthase (280–300 kDa), the chlorophyll containing monomeric PSII core complex (280 kDa) and Cyt. b_6/f complex at a molecular mass of 250–270 kDa. At about 195–210 kDa a band is present which can be visualized by Silver-stain which contains monomeric PSII core complex lacking some subunits. Two clear green bands are observed at about 70 and 140 kDa, which most likely represent the monomeric and trimeric forms of LHCII of PSII. Interestingly light harvesting proteins are prevalently present in trimeric form in dicots with respect to monocots. Finally at about 100–110 kDa a faint band containing dimeric LHCI was observed.

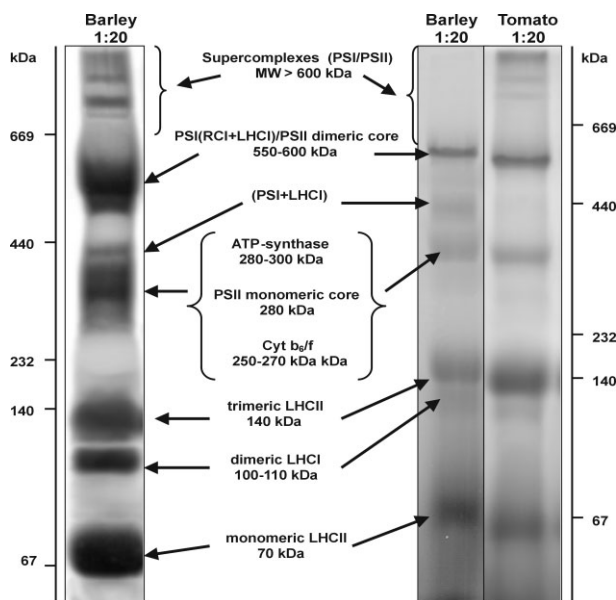


Figure 1. Blue Native electrophoresis of thylakoid membrane from barley and tomato. Thylakoid membranes were incubated with *n*-dodecyl β -D-maltoside at a 1:20 Chl/DM ratio. The gels were run according to the protocol given in section 2.2. A BNG of barley after Silver-staining is also shown on the left.

3.2 Second dimension: SDS-PAGE

All protein complexes separated by the first native dimension were directly transferred and analyzed in the second dimension by denaturing SDS-PAGE [32]. An almost regular distribution of proteins derived from complexes and supercomplexes from the first dimension was revealed where protein subunits of the thylakoid membranes were distributed by their distinct molecular masses. Specifically, a protein pattern in the 2-D gel was seen in which it was possible to recognize vertically separated lanes. Staining of acrylamide gels after 2-D-BNG/SDS-PAGE by Silver-stain revealed more spots in comparison to both coomassie R-250 and colloidal Coomassie staining (data not shown).

Figure 2 compares 2-D maps obtained starting from thylakoid membranes of tomato and barley solubilized at different Chl/DM ratios. It can be observed that at lower Chl/DM solubilization the intensity of bands in the first dimension as well as all spots in the second dimension increased, in particular those representing supercomplexes (Fig. 2B and D). This suggests that at lower Chl/DM solubilization the supercomplexes are not disassembled but there is a more

efficient extraction of all photosynthetic apparatus components from thylakoid membranes. This evidence allows one to suppose that the difference in supercomplexes observed between monocots and dicots is not related to their different solubility in detergent. On the other hand, a higher detergent concentration increased the presence of proteins localized close to the buffer front of the first dimension giving rise to a further lane in the second dimension (see arrow). This lane collects hydrophilic extrinsic proteins, like oxygen evolving proteins of PSII, and other proteins with no transmembrane domains that upon strong solubilization do not remain tightly bound to supercomplexes. 2-D maps obtained with higher Chl/DM ratios were chosen for MS analysis because of a better resolution of single spots in the gel which could reduce the possibility of protein-protein contamination. Interestingly, using the same solubilization condition (Chl/DM ratio) different species (Figs. 2B and C) gave similar patterns in the second dimension.

3.3 Mass spectrometric analysis

In order to identify subunits related to protein complexes, visible spots of membrane proteins were cut out of the SDS-PAGE, digested with trypsin and analyzed either by MALDI-TOF or by internal peptide sequencing, using mainly RP-HPLC-ESI-MS/MS. Tryptic digestion of proteins leads to peptides of different length and abundance, depending on distribution of corresponding tryptic cleavage sites (arginine and lysine). Many proteins were identified by PMF using MALDI-TOF. This technique allowed us to perform identifications of some proteins. It failed in those cases where a low number of peptides were obtained or when information on protein sequences was not present in databases. We overcame this limitation by performing internal peptide sequencing by LC coupled on-line with ESI-MS/MS. The instrument was set to prefer double charged ions, although peptides with other charge states were also considered. Moreover double charged peptides smaller than 0.6 kDa (300 m/z ; about five amino acids) were excluded with a scan range of 300–1800 m/z . Considering that the detectability decreased with increasing number of amino acids, peptides larger than 16–17 amino acids (> 1.8 kDa), were analyzed only if they contained multiple charges. In addition, peptides of more than 20 amino acids were rarely detected, especially those peptides belonging to the transmembrane domains consisting of at least 20 amino acids. Most likely these hydrophobic peptides had low charge states because of a lack of basic amino acid residues. Thus it was only possible to detect a limited number of peptides from these highly hydrophobic proteins compared to the number of peptides predicted from *in silico* analysis.

In most cases a complete amino acid sequence of a peptide was not obtained. Figure 3 shows the difference of the base peak chromatogram obtained after digestion of the spot corresponding to the protein PsbZ first with trypsin (Fig. 3A) and then by partial acid hydrolysis (Fig. 3B). In the latter case

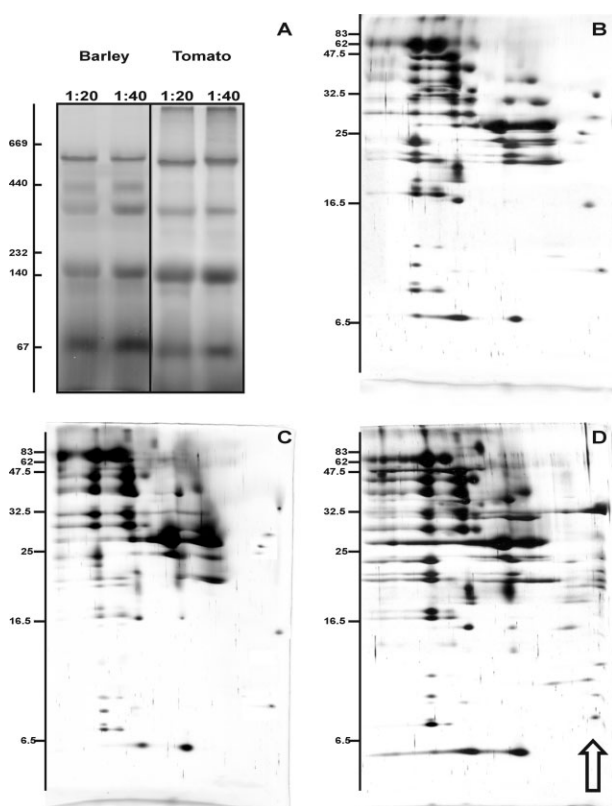


Figure 2. Comparison of BNG-PAGE and 2-D maps of barley and tomato at different DM concentrations. (A) BNG of thylakoid membranes from barley and tomato at different Chl/DM ratios (1:20, 1:40). The second dimension was performed according to Laemmli [32] (15% acrylamide, 6 M urea). (B, D) 2-D-BNG/SDS-PAGE of barley thylakoids at Chl/DM ratios of 1:20 (B) and 1:40 (D). (C) Tomato thylakoids at a Chl/DM ratio of 1:20.

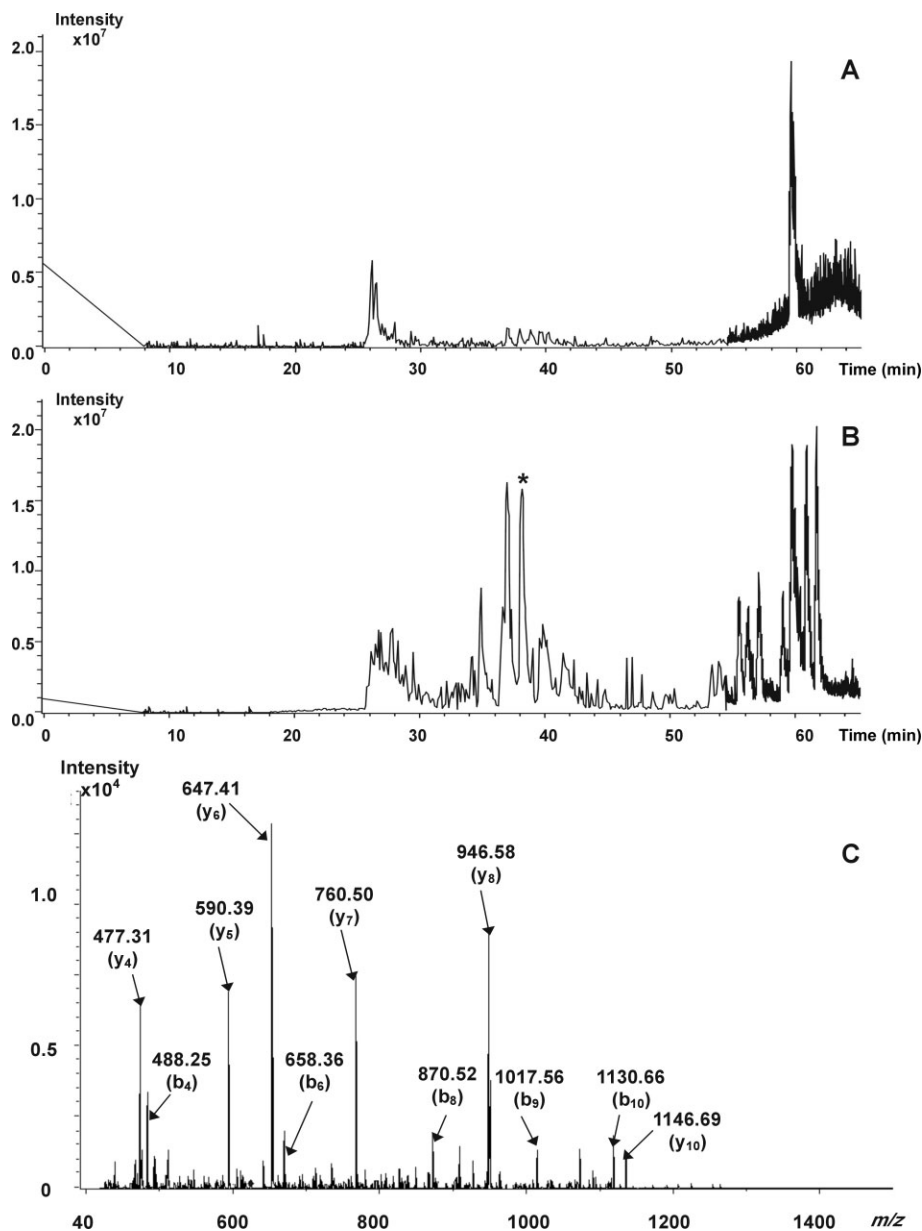


Figure 3. (A) Comparison of base peak chromatograms obtained after digestion of the gel spot corresponding to PsbZ protein by trypsin. (B) Limited acid hydrolysis of the same gel spot using 6 M HCl. (C) ESI-MS/MS spectrum of the double charged peptide ion at m/z 624.73 (TSLWIGLVFLV) marked with a star in (B).

some peptides were obtained after digestion allowing us to reach a correct identification. Figure 3C shows the tandem mass spectra acquired from the peptide ion with m/z 624.73 belonging to PsbZ and highlighted by a star in Fig. 3B. The peptide mixture was derived by partial acid hydrolysis of a single spot from the gel. In order to improve the score during data base searching, possible desamidation of glutamine and asparagine amino acids residues, were taken into account, in agreement with previous reports [36].

Tables 1–4 report the number of peptides identified and sequenced by ESI-MS/MS (first columns). The second columns show a list of peptides observed by MALDI MS. The third columns show the accession number of the identified protein and the database in which the protein was found. If the sequence(s) did not match a protein from *H. vulgare*, a homologue organism is indicated. The fourth columns show the protein M_r calculated from the amino acid sequence in the ExPASy database, whereas the apparent M_r are reported in the fifth columns according to their migration in SDS-PAGE. In the sixth columns grand average of hydropathicity (GRAVY) scores are reported and finally in the seventh columns the list of identified proteins is given (supplementary tables containing all peptide sequences obtained for each protein are available online).

In general it was observed that the total number of proteins identified by internal peptide sequencing, obtained by ESI-MS/MS, was superior to that by MALDI. Moreover, a careful analysis of peptides sequenced revealed that most of them were positioned at the soluble intermembrane

regions and were thus easily accessible for digestion with trypsin, whereas only limited information was obtained from transmembrane domains, with the exception of those proteins where a tryptic cleavage site was present inside the transmembrane domain (e.g. Psak, Psag, psbD and Lhcb1–5). However, the tryptic cleavage sites within transmembrane domains were missing from most proteins with a result of formation of oversized peptides (2–3 kDa), which escaped MS detection.

3.4 Protein identification

Figure 4 shows all the spots identified in the SDS-PAGE. The first two lanes on the left contain PSII core proteins and some of the PSII light harvesting proteins, confirming that

Table 1. Proteins identified in PSII of barley

Number of peptides identified by ESI-MS/MS	Number of peptides identified by MALDI TOF	Accession number/database (homolog organism)	Theoretical M_r	Apparent M_r	GRAVY	Protein
6	5	gi 7443223 NCBIInr Q40065 TrEMBL	25 049.6 ^{a)}	28 200	0.012	Lhcb1(.2)
2	0	gi 452341 NCBIInr (<i>Zea mays</i>) Q41823 TrEMBL	24 576.0 ^{a)}	28 200	−0.025	Lhcb2
7	6	gi 115793 NCBIInr P27523 Swiss-Prot	24 285.5	25 700	0.035	Lhcb3
7	5	gi 445116 NCBIInr	27 374.0 ^{a)}	33 600	−0.149	Lhcb4
9	7	gi 100573 NCBIInr Q40039 TrEMBL	27 137.0 ^{a)}	28 800	−0.045	Lhcb5
2	0	gi 115834 NCBIInr (<i>Lycopersicon esculentum</i>) P27525 Swiss-Prot	22 800.0 ^{a)}	22 600	0.038	Lhcb6
2	0	gi 15408693 NCBIInr (<i>Oryza sativa</i>) Q943K1 TrEMBL	21 732.1	23 900	0.433	PsbS
9	6	gi 131248 NCBIInr P05337 Swiss-Prot	38 904.6	31 000	0.373	PsbA
10	8	gi 131267 NCBIInr P10900 Swiss-Prot	56 149.1	51 000	0.088	PsbB
10	7	gi 131281 NCBIInr P11095 Swiss-Prot	52 138.1	45 200	0.264	PsbC
4	4	gi 82445 NCBIInr P11849 Swiss-Prot	39 669.5	32 100	0.319	PsbD
4	3	gi 131311 NCBIInr (<i>Triticum aestivum</i>) P05169 Swiss-Prot	9 313.4	8 000	−0.004	PsbE
1	0	gi 38605654 NCBIInr P60126 Swiss-Prot	4 367.1	4 500	0.746	PsbF
2	0	gi 38605670 P60143 Swiss-Prot	4 497.1	4 500	0.258	PsbL ^{b)}
4	4	gi 131394 NCBIInr (<i>Triticum aestivum</i>) Q00434 Swiss-Prot	20 020.3	24 100	−0.441	PsbP
7	6	gi 34914480 NCBIInr (<i>Oryza sativa</i>) Q943W1 TrEMBL	34 861.4	36 300	−0.216	PsbO
3	0	gi 140319 P08890 Swiss-Prot	6 553.7	8 100	1.376	PsbZ ^{b)}

Theoretical M_r and GRAVY scores were calculated using the Prot Param tool available at <http://us.expasy.org/>.

a) M_r reported according to Zolla *et al.* [12]

b) Protein identified by limited acid hydrolysis

the heaviest bands observed in the first dimension are prevalently supercomplexes of PSII. This evidence is in agreement with previous studies based on ultracentrifugation [39]. The third lane contains the PSI components and the dimeric core proteins from PSII. In the case of PSI, almost all proteins were identified, whereas some small proteins of PSII remained undetected. However, D1, D2, CP43 and CP47 were well identified. The identification of proteins present in

the fourth lane confirms the existence of PSI/LHCI intermediate complexes containing PsA, PsB, Lhca1, PsK, PsL and PsH. Finally the fifth lane contains a mixture of ATP-synthase complex, the monomeric core of PSII and Cyt. b_6/f .

An estimate of the quantitative ratio between the dimeric and monomeric form of the PSII core was performed using a densitometer to analyze the four most intense spots (PsbA,

Table 2. Proteins identified in PSI of barley

Number of peptides identified by ESI-MS/MS	Number of peptides identified by MALDI TOF	Accession number/database (homolog organism)	Theoretical M_r	Apparent M_r	GRAVY	Protein
7	5	gi 6685157 NCBIInr Q9SDM1 TrEMBL	22 117.2	23 400	−0.239	Lhca1
3	3	gi 1076724 NCBIInr Q43485 TrEMBL	23 161.4 ^{a)}	22 300	−0.091	Lhca2
4	4	gi 309673 NCBIInr (<i>Pisum sativum</i>) Q32904 TrEMBL	25 297.0 ^{a)}	24 600	−0.076	Lhca3
2	0	gi 9624495 NCBIInr Q9LKC7 TrEMBL	22 302.3	22 300	−0.141	Lhca4
12	10	gi 14017572 NCBIInr (<i>Triticum aestivum</i>) P58311 Swiss-Prot	83 146.6	82 500	0.252	PsaA
8	6	gi 14017571 NCBIInr (<i>Triticum aestivum</i>) P58386 Swiss-Prot	82 585.3	82 500	0.104	PsaB
1	0	gi 100633 NCBIInr P10794 Swiss-Prot	8 768.1	7 900	−0.080	PsaC
7	6	gi 548603 NCBIInr P36213 Swiss-Prot	17 559.0	21 800	−0.398	PsaD
4	0	gi 131176 NCBIInr P13194 Swiss-Prot	10 821.1	18 500	−0.658	PsaE
5	5	gi 548604 NCBIInr P13192 Swiss-Prot	17 464.2	18 100	−0.040	PsaF
3	3	gi 131192 NCBIInr Q00327 Swiss-Prot	10 810.2	10 200	0.093	PsaG
2	0	gi 131196 NCBIInr P20143 Swiss-Prot	10 193.6	12 300	−0.187	PsaH
2	0	gi 131211 (<i>Triticum aestivum</i>) P19443 Swiss-Prot	4 744.6	7 200	−0.879	PsaJ ^{b)}
1	0	gi 131225 NCBIInr P23993 Swiss-Prot	17 987.7	17 700	0.331	PsaL
1	0	gi 400879 NCBIInr P31093 Swiss-Prot	9 817.1	9 400	−0.699	PsaN
3	0	gi 548605 NCBIInr P36886 Swiss-Prot	9 022.4	7 200	0.461	PsaK

Theoretical M_r and GRAVY scores were calculated using the Prot Param tool available at <http://us.expasy.org/>.

a) M_r reported according to Zolla *et al.* [11]

b) Protein identified by limited acid hydrolysis

PsbB, PsbC, PsbD) in the colloidal coomassie stained gels. This analysis gave a ratio of 2:1 which is not in agreement with what is commonly observed by ultracentrifugation, where the ratio is 1.3:2 [2]. This indicates that under the experimental condition used in this study a possible aggregation of monomers is induced. The sixth SDS lane shows the existence of a PSII core pseudo-complex, lacking PsbC and PsbF proteins. The PSII antenna proteins were detected both in supercomplexes and in the seventh and ninth lanes, corresponding to the monomeric and trimeric state of PSII light harvesting proteins. As expected, the minor antenna

Lhcb4, Lhcb5 and Lhcb6 are present in a monomeric state, while the major antenna Lhcb1, Lhcb2 and Lhcb3 are either in a monomeric or prevalently in a trimeric state [3]. Interestingly the band in the first dimension just below the trimeric LHCII, contains PSI antenna, which before adding SDS were in dimeric aggregation, as suggested by many functional studies [40].

Finally, the buffer front of native PAGE represented the more hydrophilic proteins, like the PSII oxygen-evolving complex, or proteins containing no transmembrane domain(s) which escaped from the complexes during solu-

Table 3. Proteins identified in Cyt. *b₆/f* of barley

Number of peptides identified by ESI-MS/MS	Number of peptide identified by MALDI TOF	Accession number/database (homolog organism)	Theoretical M_r	Apparent M_r	GRAVY	Protein
13	16	gi 14017584 (<i>Triticum aestivum</i>) P05151 Swiss-Prot	31 304.2	37 800	−0.142	PetA
5	4	gi 82433 P60161 Swiss-Prot	26 093.8	20 200	0.522	PetB
5	5	gi 32394644 Q7X9A6 TrEMBL	23 726.2	21 400	−0.069	PetC
2	0	gi 82437 P12361 Swiss-Prot	17 516.8 ^{a)} (<i>Triticum aestivum</i> P12119)	16 900	0.566	PetD
2	0	gi 130269 P08248 Swiss-Prot	10 184.3	31 700	−0.006	PetE
4	0	gi 20302473 (<i>Triticum aestivum</i>) Q8RVZ8 Swiss-Prot	40 232.0		−0.410	
				38 900		
3	0	gi 20302471 (<i>Triticum aestivum</i>) Q8RVZ9 Swiss-Prot	38 806.5		−0.312	PetH

Theoretical M_r and GRAVY scores were calculated using the Prot Param tool available at <http://us.expasy.org/>.

a) M_r of the complete protein chain of *Triticum aestivum* was reported because only a fragment of the corresponding protein of barley was present in the databases used

Table 4. Proteins identified in ATP synthase of barley

Number of peptides identified by ESI-MS/MS	Number of peptides identified by MALDI TOF	Accession number/database (homolog organism)	Theoretical M_r	Apparent M_r	GRAVY	Protein
23	21	gi 14017569 NCBIInr (<i>Triticum aestivum</i>) P12112 Swiss-Prot	55 295.5	57 500	−0.068	AtpA
17	15	gi 114548 NCBIInr P00828 Swiss-Prot	53 875.5	57 500	−0.094	AtpB
4	4	gi 27818005 NCBIInr <i>Oryza sativa</i> (japonica cultivar-group) Q84NW1 TrEMBL	39 706.8	36 300	−0.091	AtpC
9	8	gi 67870 NCBIInr P20859 Swiss-Prot	15 189.6	19 000	0.011	AtpE
2 ^{a)}	0	gi 11466782 NCBIInr (<i>Oryza sativa</i>) P00843 Swiss-Prot	7 974.4	4 200	1.035	AtpH

Theoretical M_r and GRAVY scores were calculated using the Prot Param tool available at <http://us.expasy.org/>.

a) Nontryptic fragments

bilization. The latter, derived from PSI, like the gene products of PsaN (GRAVY: −0.699) or gene products PsbO and PsbP (GRAVY: −0.216 and −0.441). In contrast, gene products PsuC, PsuD and PsuE from PSI, which also lack the transmembrane domains, remained bound to the PSI complex.

Thus, most spots were identified and most expected proteins were found, with the exception of small proteins (under 8–10 kDa) of which only PsbF (4.4 kDa) was identified. It is noticeable that some small proteins, like PsbZ, PsbL and Psaj, which were hardly detected using tryptic digestion,

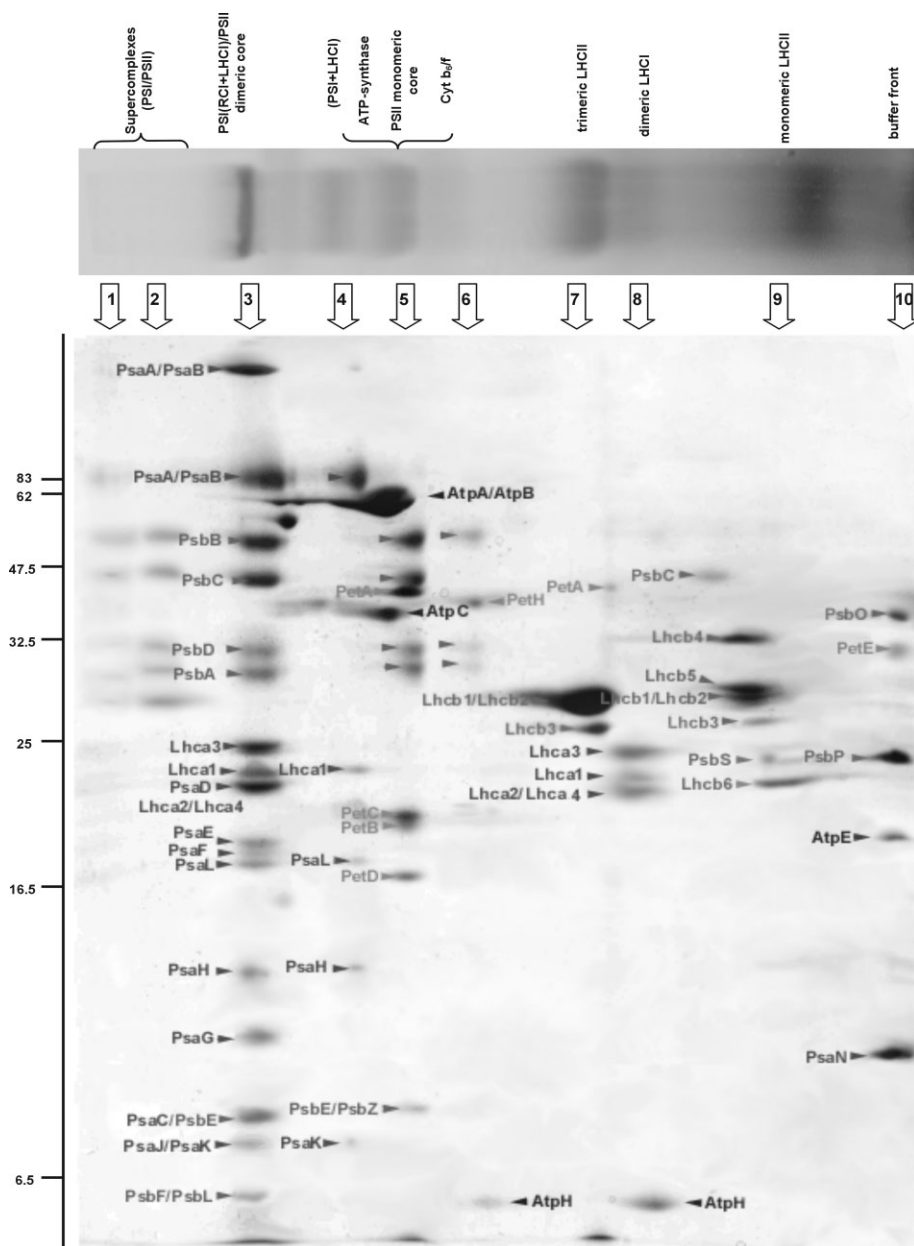


Figure 4. 2-D-BNG/SDS map of barley thylakoid membranes showing proteins identified by ESI-MS/MS and MALDI. In the first dimension thylakoid membranes were incubated with DM at a 1:20 Chl/DM ratio. In the second dimension a solubilized lane was loaded onto a Laemmli gel (15% acrylamide, 6 M urea). Numbers inside arrows indicate vertical disposition in lanes of subunits belonging to each complex visible in the map. Different colors were used to identify proteins belonging to different complexes: red (PSI), green (PSII), blue (ATP-synthase) and orange (Cyt. *b₆/f*).

could be identified upon cleavage by limited acidic hydrolysis. Examination of the sequences of these proteins revealed that they were not likely to produce tryptic fragments which could be observed in the mass range investigated by ESI MS.

4 Discussion

The protocol for thylakoid membrane proteins provided here, allowed to a 2-D differential gel display of most hydrophilic and hydrophobic proteins of the thylakoid membrane to be performed. The BNG/SDS-PAGE technique provided an excellent alternative to 2-D IEF/SDS-PAGE for 2-D separation of the most hydrophobic thylakoid proteins. IEF in the first dimension excludes some integral membrane proteins as a result of poor solubility in the neutral or zwitterionic detergents that must be used in IEF [41, 42]. Although addition of amidosulfobetaine detergents along with thiourea [43–45] as well as addition of DM [46] have increased protein separation in some cases, attempts to use this approach for a complete proteomic analysis of thylakoids have not been productive, to date.

Hippler and coworkers [19] reported the presence of highly hydrophobic PsaA in 2-D IEF/SDS-PAGE gels of thylakoid membrane from *Chlamydomonas* using DM, giving confidence that the larger integral proteins may be mapped, but most proteins from the PSII core complex were not revealed. The efficient protein transfer obtained by 2-D-BNG/SDS-PAGE allows a continuous 2-D with respect to IEF/SDS-PAGE minimizing any protein loss. In fact, no further fixation or staining of proteins is necessary after the first dimension, therefore proteins are denatured directly within the gel and then transferred completely into the second dimension SDS-PAGE. However, the biggest advantage of BNG is that addition of a modest amount of Coomassie Blue dye maintains the integrity of integral protein complexes and consequently the subsequent second dimension gel gives information on the composition of protein complexes. In addition, certain proteins can be observed to be more loosely bound to the complexes due to their presence in different lanes in the second dimension.

Significant differences in protein aggregation states are observed between monocots and dicots, as the latter phyla contain more stable PSII supercomplexes. Moreover trimeric forms of antenna are more prevalent in dicots than in monocots. This is in agreement with our previous results indicating generally less heterogeneity of Lhcb1 in monocots than in dicots [10] and with a previous report that dicots appeared to generate higher ratios of oligomeric to monomeric antenna system *in vitro* [47]. Our data show that trypsin digestion produces only a limited number of peptides from these highly hydrophobic proteins as compared to peptides predicted from *in silico* analysis, although strong denaturing conditions were used. This probably reflects the fact that tryptic cleavage is disfavored in the hydrophobic regions. From a statistical analysis of the total number of peptides obtained and the average sequence length *per* cleavage site, it clearly appeared that more hydrophilic proteins generated more peptides with an average tryptic cleavage distance of about 14–15 amino acids promoting their detectability by MS, whereas the more hydrophobic proteins lack a sufficient number of tryptic cleavage sites within transmembrane domains, limiting their detection by ESI-MS/MS to a mass range between 300 to 1800 *m/z*. In some cases, however, the transmembrane domain of membrane proteins included tryptic cleavage sites, which facilitated their detection and identification (*e.g.* PsaK, PsaG, PsbD and Lhcb1–5). Consequently, in many cases the limited number of peptides prevented unequivocal protein identification from PMF, particularly when using MALDI-TOF and identification was mostly probable by sequencing of peptides from the hydrophilic regions.

Most of the protein subunits from PSI (about 90%) were detected, whereas in the case of PSII only about 50% were detected. The low molecular mass proteins (smaller than 10 kDa) of PSII escaped identification. However, cleavage by acid hydrolysis allowed identification of some of the small proteins, such as PsbZ, PsbL and PsaJ, indicating that small protein size is not the limiting factor using SDS-PAGE and that the lack of internal tryptic-cleavage sites or their low abundance may have prevented their identification by standard tryptic digestion. Proteins localized at the buffer front during native PAGE were released during BNG-PAGE. They represent the more hydrophilic proteins, or proteins containing no transmembrane domains, which escaped from the complexes during solubilization. The gene products PsaN from PSI as well as PsbO and PsbP from the oxygen-evolving complex of PSII extrinsic phases were found in the buffer front. Interestingly, PsaC, PsaD and PsaE from PSI which also contain no transmembrane domain remained bound to the PSI complex, indicating a strong interaction between PSI and these subunits.

From data reported in this study, it can be asserted that the 2-D native/SDS-PAGE system seems particularly well suited for separation of highly hydrophobic membrane proteins with GRAVY scores higher than 0.5, such as PetB and

PetD (GRAVY: 0.522 and 0.566, respectively), PSII core protein PsbF and PsbZ (GRAVY: 0.746 and 1.376, respectively) and AtpH (GRAVY score: 1.035), as well as proteins with high *pI* values that fail to enter the first dimension of IEF gels [48], allowing one to conclude that the coverage of integral thylakoid membranes by 2-D BNG/SDS-PAGE was satisfactory. In the case of smaller proteins the separation method used is also not the limiting factor. In fact, although, BNG electrophoresis coupled with SDS-PAGE, is a method which lacks true orthogonality, since separation in the two dimensions is driven by mass-discrimination, small proteins are easily identified using limited acid hydrolysis. This is in agreement with a previous report demonstrating that very hydrophobic proteins, known to withstand treatment by trypsin, may be included in the scope of proteomic research [36]. However, in the particular case of core proteins, the low abundance of some of these small proteins might have resulted in their lack of detection.

The protein molar ratio within proteomes varies over several orders of magnitude and in many cases only the most abundant proteins are identified. We are therefore working to develop a novel prefractionation procedure which may guarantee an enrichment of low abundant proteins before 2-D analysis. Finally, further progress must be made for analysis of isomeric forms and PTMs such as oxidation, *etc.* In this regard the chromatographic proteome of the thylakoid membrane allows an alternative strategy to a complete proteomic analysis of thylakoid membrane [49]. Recently a 2-D chromatography system, directly interfaced to ESI MS, has allowed the direct visualization of large reaction-center proteins (up to 83 kDa) from both PSI and PSII, providing an attractive avenue for characterization of thylakoid membrane proteomes [50].

5 Concluding remarks

In conclusion, proteomics still needs methodological improvements to allow full characterization of all the proteins in complex systems although the techniques described here provide high-throughput for identification of most integral membrane proteins. It is likely that a combination of 2-DE and 2-D LC, as recently reported [1], will facilitate many new insights into the thylakoid proteome, complementing the existing body of knowledge. However, the final aim should be to characterize the thylakoid proteome experimentally as much as possible, to identify the photosynthetic apparatus in its completeness and identify known and potential components expressed during the stress response, using simple and repeatable methods which avoid multiple extractions and manipulations. Only in this way, it will be possible in the near future to achieve a direct proteomic comparison of the photosynthetic apparatus under different conditions with the aim of better understanding the molecular mechanisms underlying physiological adaptations or for screening of photosynthetic mutants.

We want to thank Marjaana Suorsa and Virpi Pakkaarinen from Turku University, Shabaz Mohammed from University of Southern Denmark, Anna Maria Timperio, Gian Maria D'Amici and Sara Rinalducci from Tuscia University for their precious help. We also want to thank the CIB (Consorzio Inter-universitario per le Biotecnologie) who supported this study. Thanks to the Progetto Strategico 1999 (Area 2.2).

6 References

- [1] Friso, G., Giacomelli, L., Ytterberg, A. J., Peltier, J. B. *et al.*, *Plant Cell* 2004, **16**, 478–499.
- [2] Hankamer, B., Nield, J., Zheleva, D., Boekema, E. *et al.*, *Eur. J. Biochem.* 1997, **243**, 422–429.
- [3] Jansson, S., *Biochim. Biophys. Acta* 1994, **1184**, 1–19.
- [4] Scheller, H. V., Jensen, P. E., Haldrup, A., Lunde, C., Knoetzel, J., *Biochim. Biophys. Acta* 2001, **1507**, 41–60.
- [5] Poetsch, A., Rexroth, S., Heberle, J., Link, T. A. *et al.*, *Biochim. Biophys. Acta* 2003, **1618**, 59–66.
- [6] Whitelegge, J. P., Zhang, H., Aguilera, R., Taylor, R. M., Cramer, W. A., *Mol. Cell. Proteomics* 2002, **1**, 816–827.
- [7] Jansson, S., Pichersky, E., Bassi, R., Green, B. R. *et al.*, *Plant Mol. Biol. Rep.* 1992, **10**, 242–253.
- [8] Dainese, P., Bassi, R., *J. Biol. Chem.* 1991, **266**, 8136–8142.
- [9] Corradini, D., Huber, C. G., Timperio, A. M., Zolla, L., *J. Chromatogr. A* 2000, **886**, 111–121.
- [10] Huber, C. G., Timperio, A. M., Zolla, L., *J. Biol. Chem.* 2001, **276**, 45755–45761.
- [11] Zolla, L., Rinalducci, S., Timperio, A. M., Huber, C. G., *Plant Physiol.* 2002, **130**, 1938–1950.
- [12] Zolla, L., Timperio, A. M., Walcher, W., Huber, C. G., *Plant Physiol.* 2003, **131**, 198–214.
- [13] Gomez, S. M., Nishio, J. N., Faull, K. F., Whitelegge, J. P., *Mol. Cell. Proteomics* 2002, **1**, 46–59.
- [14] Rabilloud, T., Charmont, S., in: Rabilloud, T. (Ed.), *Proteome Research: Two-dimensional Gel Electrophoresis and Identification Methods*, Springer, Berlin, Germany 2000, pp. 107–126.
- [15] Peltier, J. B., Friso, G., Kalume, D. E., Roepstorff, P. *et al.*, *Plant Cell* 2000, **12**, 319–341.
- [16] van Wijk, K. J., *Plant Physiol.* 2001, **126**, 501–508.
- [17] Santoni, V., Rabilloud, T., Doumas, P., Rouquie, D. *et al.*, *Electrophoresis* 1999, **20**, 705–711.
- [18] Santoni, V., Kieffer, S., Desclaux, D., Masson, F., Rabilloud, T., *Electrophoresis* 2000, **21**, 3329–3344.
- [19] Hippler, M., Klein, J., Fink, A., Allinger, T., Hoerth, P., *Plant J.* 2001, **28**, 595–606.
- [20] Schagger, H., Cramer, W. A., von Jagow, G., *Anal. Biochem.* 1994, **217**, 220–230.
- [21] Devreese, B., Vanrobaeys, F., Smet, J., Van Beeumen, J., Van Coster, R., *Electrophoresis* 2002, **23**, 2525–2533.
- [22] Kuchler, M., Decker, S., Hormann, F., Soll, J., Heins, L., *EMBO J.* 2002, **21**, 6136–6145.
- [23] Thidholm, E., Lindstrom, V., Tissier, C., Robinson, C. *et al.*, *FEBS Lett.* 2002, **513**, 217–222.
- [24] Kruff, V., Eubel, H., Jansch, L., Werhahn, W., Braun, H. P., *Plant Physiol.* 2001, **127**, 1649–1710.
- [25] Herranen, M., Battchikova, N., Zhang, P., Graf, A. *et al.*, *Plant Physiol.* 2004, **134**, 470–481.
- [26] Rexroth, S., Meyer zu Tittingdorf, J. M. W., Krause, F., Dencher, N. A. *et al.*, *Electrophoresis* 2003, **24**, 2814–2823.
- [27] Heinemeyer, J., Eubel, H., Wehmhöner, D., Jansch, L. *et al.*, *Phytochemistry* 2004, **65**, 1683–1692.
- [28] Berthold, D. A., Babcock, G. T., Yocum, C. F., *FEBS Lett.* 1981, **134**, 231–234.
- [29] Zolla, L., Timperio, A. M., Testi, M. G., Bianchetti, M. *et al.*, *Photosynth. Res.* 1999, **61**, 281–290.
- [30] Kügler, M., Jansch, L., Kruff, V., Schmitz, U. K., Braun, H. P., *Photosynth. Res.* 1997, **53**, 35–44.
- [31] Suorsa, M., Regel, R. E., Paakkari, V., Battchikova, N. *et al.*, *Eur. J. Biochem.* 2004, **271**, 96–107.
- [32] Laemmli, U. K., *Nature* 1970, **227**, 680–685.
- [33] Blum, H., Beier, H., Gross, H. J., *Electrophoresis* 1987, **8**, 93–99.
- [34] Porra, R. J., Thompson, W. A., Kriedemann, P. E., *Biochim. Biophys. Acta* 1989, **975**, 384–394.
- [35] Shevchenko, A., Wilm, M., Vorm, O., Mann, M., *Anal. Chem.* 1996, **68**, 850–858.
- [36] Shevchenko, A., Loboda, A., Ens, W., Standing, K. G., *Anal. Chem.* 2000, **72**, 2132–2141.
- [37] Gobom, J., Nordhoff, E., Mirgorodskaya, E., Ekman, R., Roepstorff, P., *J. Mass Spectrom.* 1999, **34**, 105–116.
- [38] Larsen, M. R., Cordwell, S. J., Roepstorff, P., *Proteomics* 2002, **2**, 1277–1287.
- [39] Eshaghi, S., Andersson, B., Barber, J., *FEBS Lett.* 1999, **4446**, 23–26.
- [40] Jansson, S., Andersen, B., Scheller, H. V., *Plant Physiol.* 1996, **112**, 409–420.
- [41] Molloy, M. P., *Anal. Biochem.* 2000, **280**, 1–10.
- [42] Santoni, V., Molloy, M., Rabilloud, T., *Electrophoresis* 2000, **21**, 1054–1070.
- [43] Pasquali, C., Fialka, I., Huber, L. A., *Electrophoresis* 1997, **18**, 2573–2581.
- [44] Chevallet, M., Santoni, V., Poinas, A., Rouquie, D. *et al.*, *Electrophoresis* 1998, **19**, 1901–1909.
- [45] Rabilloud, T., *Methods Mol. Biol.* 1999, **112**, 9–19.
- [46] Naver, H., Boudreau, E., Rochaix, J. D., *Plant Cell* 2001, **13**, 2731–2745.
- [47] Huner, N. P. A., Campbell, D., Krol, M., Hayden, B. D. *et al.*, *Plant Physiol.* 1992, **99**, 830–836.
- [48] Rowley, A., Choudhary, J. S., Marzioch, M., Ward, M. A. *et al.*, *Methods* 2000, **20**, 383–397.
- [49] Huber, C. G., Walcher, W., Timperio, A. M., Troiani, S. *et al.*, *Proteomics* 2004, **4**, 3909–3920.
- [50] Whitelegge, J. P., *Photosynth. Res.* 2003, **78**, 265–277.

PARTE III

**“Elettroforesi bidimensionale e
spettrometria di massa
nell’identificazione di proteine
caratterizzanti le specie *Coffea arabica*
(var. Colombia) e *Coffea canephora*
(var. Indiano Robusta)”**

REGULAR ARTICLE

Two-dimensional mapping as a tool for classification of green coffee bean species

Maria Teresa Gil-Agusti¹*, Natascia Campostrini²*, Lello Zolla³, Corrado Ciambella³, Carlo Invernizzi⁴ and Pier Giorgio Righetti²

¹ Area Química Analítica, CCEE, ESTCE, Universitat Jaume I, Castellon, Spain

² Laboratory of Proteome Analysis, University of Verona, Verona, Italy

³ Department of Environmental Sciences, University of Tuscia, Viterbo, Italy

⁴ Caffè Invernizzi, Via G. Compagnoni 29, Milano, Italy

Two species of the genus *Coffea*, *Coffea arabica* (Colombia) and *Coffea canephora* (Indiano Robusta) were analysed by two-dimensional (2-D) maps in order to obtain fingerprints of the expressed polypeptide chains and to determine which ones would characterize the two species. Green beans were milled under liquid nitrogen. A dry powder was produced by three different extraction protocols aimed at eliminating interfering substances (polyphenols). A reduced powder was produced by two successive extractions performed in acetone. Trichloroacetic acid (TCA; 10% w/v) and β -mercaptoethanol (0.07% v/v) in acetone were used for the first extraction (a) and 10% w/v TCA in acetone was used for the second extraction (b). Proteins were then solubilized in a solution (40 μ L per 1 mg powder) containing 7 M urea, 2 M thiourea, 3% w/v 3-(3-cholamido-propyldimethyl-amino)-1-propanesulfate, 1% v/v carrier ampholytes, 40 mM Tris, 5 mM tributylphosphine and 10 mM acrylamide as alkylating agent. Following incubation at room temperature for 1 hour and centrifugation (7000 rpm for 20 minutes), the supernatant was used for 2-D electrophoresis. The proteins were revealed by Sypro Ruby staining. Master maps of the five replicas of each species were compared by PDQuest analysis. The results of this differential proteome analysis were: sixteen proteins were expressed solely in *C. canephora* (var. Indiano Robusta) and five proteins were only found in *C. arabica* (var. Colombia). Another eight proteins were up-regulated in *C. canephora* (var. Indiano Robusta) in comparison to *C. arabica* (var. Colombia) and one was down-regulated in the same comparison. A number of these polypeptide chains were further characterized by mass spectrometry in the matrix-assisted laser desorption/ionisation-time of flight mode. Additionally, considering the low number of protein sequences of *Coffea* present in the databases we also investigated some spots with a more powerful tool, reversed phase-high-performance liquid chromatography-electrospray ionisation-tandem mass spectrometry, thus obtaining an internal peptide sequence. The general properties of the identified proteins are presented and discussed.

Received: July 14, 2004
Revised: September 9, 2004
Accepted: September 10, 2004

Keywords:

Green coffee beans / Two-dimensional maps

Correspondence: Professor Pier Giorgio Righetti, Department of Agricultural and Industrial Biotechnologies, University of Verona, Strada Le Grazie 15, 37134 Verona, Italy
E-mail: righetti@sci.univr.it
Fax: +39-045-8027929

Abbreviations: AA, acrylamide; AUC, Acetic acid-Urea-CTAB; CTAB, cetyltrimethylammonium bromide; FA, formic acid; TBP, tributylphosphine

1 Introduction

Ever since its introduction in Europe in the second half of the sixteen century [1] (probably spreading out from Venice, in those days aptly considered the emporium of Orient), coffee beans have been steadily growing in importance, due to their

* These authors contributed equally.

capillary diffusion first as an expensive beverage in coffee houses and then as a common household item. Today, this crop is believed to be one of the most important primary materials in the world, in terms of revenue generated, coming second only to petrol, in this respect [2]. The > 150 varieties grown around the world can be grouped into two species of the genus *Coffea* (family Rubiaceae), *Coffea arabica* (var. Colombia) and *Coffea canephora* (var. Indiano Robusta). Since the beginning of the sixties, with the advent of GC and MS, coffee has been extensively studied, with the aim in particular of isolating and characterizing the substances which impart to it its unique flavor. To date, > 850 compounds have been identified in the flavor of roasted coffee, and > 300 in the smell of green coffee [3], no doubt also due to the advent of the 2-D technique GC $\times$ GC, which has greatly helped in isolating minor species hidden in major peaks [4].

However, very few studies have investigated the protein content of coffee beans, especially in regard to the use of 2-D maps, although it is well known that free amino acids, peptides and proteins are essential aroma precursors [5]. In this respect, perhaps the cornerstone study is the one by Rogers *et al.* [6], who reported the first detailed investigation of proteins in the *C. arabica* endosperm by 2-D map analysis, coupled to MS characterization of the most abundant polypeptide chains and mRNA amplification. In their study, the most abundant polypeptide spots observed on mature coffee grain 2-D profiles were found to be subunits of the same protein, which exist as multiple isoforms with varying pI values. Strong sequence similarities were found to the 11S family of plant storage proteins. Their structure is typical of the 11S type, which occurs as a precursor of a 55 kDa polypeptide, and is observed on 2-D profiles, under denaturing and reducing conditions, in the form of cleavage products at approximately 22 kDa (β arm) and 33 kDa (α arm). This 11S-like-protein accounts for about 45% of the total grain protein.

In another study, Ludwig *et al.* [7] investigated peptides and proteases in green coffee beans of both *C. arabica* and *C. canephora* by 2-D maps and *in situ* zymogramming. They were able to detect at least five to seven enzyme spots in all coffee samples tested, whose activities were stable for several months in intact beans. Moreover, by using Schagger and von Jagow's buffer system [8] in the second dimension gel, they could resolve peptides typically lost in conventional 2-D maps in their gel slabs which ranged in size from as low as 3500 Da up to 10 000 Da. By a model roasting carried out with these coffee peptides, as well as with peptide-dextrose mixtures, they could confirm the assumption that peptides contribute to the roasting flavor of coffee and build up a map of intensities of sensory impression.

In another line of research, Montavon *et al.* [9, 10] studied the evolution of green coffee protein profiles upon maturation and the changes in these protein profiles upon roasting, both *via* 2-D map analysis. In the first study [9], they confirmed, in both ripe and unripe samples, the presence of the α and β arms, as subunits of the 11S-like protein, as found by

Ludwig *et al.* [7], with no marked differences in the 2-D profiles between the two classes of beans. However, compared to ripe beans, unripe specimens were found to be more sensitive to oxidation of chlorogenic acids to such an extent that even upon incubation under anaerobic conditions, massive losses of α and β arm polypeptides could be detected in 2-D maps, with production of low M_r peptides. When investigating the changes in protein profiles upon roasting [10], however, not much could be deduced from 2-D map analyses, since, when reaching the high temperatures of a proper roasting process (240°C), essentially all polypeptide chains disappeared from the 2-D maps. In these studies no attempts were made to use 2-D maps for phylogenetic studies, *i.e.* for characterizing different coffee species *via* differential, quantitative proteomic analysis to classify the different varieties *via* polypeptide chains present in only single species and/or up- or down-regulated in the various cultivars. We report here the first such attempt, *via* a preliminary study on two species of the genus *Coffea*, *C. arabica* (var. Colombia) and *C. canephora* (var. Indiano Robusta).

2 Materials and methods

2.1 Chemicals and materials

Tris, mineral oil, DL-DTT, Tween 20, cetyltrimethylammonium bromide (CTAB) and EDTA were purchased from Sigma (Steinheim, Germany). Glycerol, methanol, ethanol, acetic acid and acetone were from Merck (Darmstadt, Germany). Acrylamide/Bis solution (40%), TEMED, acrylamide (AA), ammonium persulfate, the Protean IEF Cell, the GS-710 Densitometer, the Versa Doc Scanner, PDQuest Version 6.2 software as well as the linear Immobiline dry strips pH gradient 3–10 (17 cm long), were obtained from Bio-Rad (Hercules, CA, USA); glycine, SDS, iodoacetamide, urea, thiourea, tributylphosphine (TBP) and CHAPS were from Fluka (Buchs, Switzerland). Bromophenol blue, carrier ampholytes and agarose were purchased from Pharmacia Diagnostics (Uppsala, Sweden).

2.2 2-DE

2.2.1 Plant material

Indiano Robusta green coffee samples (*C. canephora* var. Indiano Robusta type Cherry AB) were obtained from Western India. Cherries were harvested in January 2002 and exported to Italy in March 2002. The beans were picked at full ripening (*ca.* 9–11 months on the plant). After removing the bean from the epicarp, it was subjected to natural dessication (sun-dried) for *ca.* 12–15 days until its humidity reached 12%. Colombia green coffee samples (*C. arabica* var. Colombia type Supreme) were obtained from Medellin, Western Colombia. Cherries were harvested from the plantation on April 2002 and exported to Italy in June 2002. The

beans were picked at full ripening (*ca.* 6–8 months on the plant). After harvesting, the bean was removed from the epicarp by a machine treatment. After washing, the mucilage was removed by enzyme action. A subsequent washing step removed the enzymes and the pergamin, a thin skin covering the bean. Finally, the beans were dried. These samples were kindly provided by C. Invernizzi (Casa del Caffè, Milano, Italy). Green beans from *C. arabica* and *C. canephora*, divided in two classes, constituted the dataset used in this study: (i) green beans from *C. arabica* (var. Colombia; the pool used to perform six 2-D maps; Fig. 1a); (ii) green beans from *C. canephora* (var. Indiano Robusta; the pool used to perform six 2-D maps; Fig. 1b). The procedure was repeated three times, thus producing three genuine replicates for each map.

2.2.2 Protein extraction and preparation of samples for electrophoresis

Grains were finely ground in liquid nitrogen. A dry powder was produced by three different extraction protocols aimed at eliminating interfering substances (such as polyphenols).

2.2.2.1 Acetone/TCA/ β -mercaptoethanol solution extraction

A dry, reduced powder was produced by two successive extractions performed in acetone. TCA (10% w/v) and β -mercaptoethanol (0.07% v/v) in acetone were used for the first extraction (a) and 10% w/v TCA in acetone was used for the second extraction (b), according to the method of Damerval *et al.* [11]. Proteins were then solubilized in a solution (40 μ L per 1 mg powder) containing 7 M urea, 2 M thiourea, 3% w/v CHAPS, 1% v/v carrier ampholytes (pH 3–10), 40 mM Tris, 5 mM TBP, 10 mM AA [12]. Following incubation at room temperature for 1 h (alkylation was blocked with 10 mM DTT) and centrifugation (7000 rpm for 20 min, 4°C) the supernatant was stored at –20°C until use.

2.2.2.2 Lysis solution extraction

Green coffee samples were milled under nitrogen and extracted in solubilization/lysis buffer (1 mL per 200 mg powder) containing 7 M urea, 2 M thiourea, 3% w/v CHAPS, 1% v/v carrier ampholytes (pH 3–10), 40 mM Tris, 5 mM TBP, 10 mM AA [12]. Extraction was allowed to proceed for 1 h at room temperature under continuous shaking and alkylation was blocked with 10 mM DTT. Extracts were then centrifuged (7000 rpm, 20 min, 4°C) and supernatants were stored at –20°C until use.

2.2.2.3 Acetic acid/Urea/CTAB solution extraction

Green coffee samples were milled under nitrogen and extracted in a solution (1 mL per 200 mg powder) containing 0.1 M acetic acid, 3 M urea, 0.01% CTAB. Extraction was

allowed to proceed for 2 h at room temperature with continuous shaking. Extracts were then centrifuged (13 000 rpm, 25 min, 4°C) and supernatants were precipitated in an anhydrous solution of acetone and methanol in an 8:1 v/v ratio. The samples were kept at –20°C for 2 h and then centrifuged at 13 000 rpm for 30 min. The supernatant was discarded and the pellet was resuspended in a solution (40 μ L per 1 mg powder) containing 7 M urea, 2 M thiourea, 3% w/v CHAPS, 1% v/v carrier ampholytes (pH 3–10), 40 mM Tris, 5 mM TBP, 10 mM AA. Following incubation at room temperature for 1 h (alkylation was blocked with 10 mM DTT) and centrifugation (7000 rpm for 20 min, 4°C) the supernatant was stored at –20°C until use. We refer to this method as the acetic acid/urea/CTAB (AUC) extraction method.

2.2.3 IEF in IPG strips

The first dimension run was performed on strips (17 cm length, 0.5 mm thickness) with a linear pH gradient from pH 3 to pH 10 [13]. For analytical runs (used to find differentially expressed proteins), the IPG strips were rehydrated with 3 mg/mL of pooled samples (six strips with *C. arabica* samples and six strips with *C. canephora* samples) and traces of bromophenol blue for monitoring the electrophoretic run. For preparative runs (used to isolate proteins from the gel for identification with MS), the IPG strips were rehydrated with 4.5 mg/mL of pooled samples. Passive gel rehydration was allowed to continue for 8 h before the focusing step. IEF was carried out with a Protean IEF Cell (Bio-Rad), with a low initial voltage and then by applying a voltage gradient up to 10 000 V with a limiting current of 50 μ A/strip. The total product time $\times$ voltage applied was 75 000 Vh for each strip and the temperature was set at 20°C.

2.2.4 Interfacing the IPG strips with the denaturing SDS solution and SDS-PAGE

Each strip was equilibrated with an SDS denaturing solution containing 6 M urea, 2% SDS, 20% glycerol, 0.375 M Tris-HCl (pH 8.8). The contact lasted for 30 min in tubes containing 20 mL each of the equilibration solution. Each strip was then interfaced with a gel slab using 0.5% agarose solubilized in cathodic buffer (192 mM glycine, 0.1% SDS, Tris to pH 8.3). The gel slabs were cast with a two-vessel gradient mixer, with total sufficient volumes for polymerizing 12 gel slabs of 1.5 mm thickness and with a porosity gradient from 7%T to 20%T. Polymerization took place overnight. The second dimension run was performed using a PROTEAN II xl Multi-Cell (Bio-Rad). The cathodic buffer was the same as stated above; the anodic buffer was a solution of 0.375 M Tris-HCl, pH 8.8. The electrophoretic run was performed by setting a current of 2 mA for each gel for 1 h, then 5 mA/gel for 2 h and 10 mA/gel until the end of the run for a total of 20–22 h. During the whole run the temperature was kept constant at 11°C.

2.3 Fluorescent staining

All gels were fixed with a solution containing 40% ethanol and 10% acetic acid for 30 min. After this step the gels were stained with Sypro Ruby (70 mL for each gel) overnight and then destained with 10% methanol, 7% acetic acid for 1 h [14]. The gels were then washed with milliQ water (Millipore, Bedford, MA, USA). Samples were digitized with a Versa Doc Scanner. The gels used to isolate proteins for identification with MS were stained with colloidal Coomassie blue.

2.4 Protein pattern and statistical analysis

Digitized images were acquired with PDQuest software (version 6.2) [15], which was used for cropping and orienting the images, detecting and identifying spots, comparing and matching spots, normalizing and analyzing the data and for preparing a report. A match set was created from the protein patterns of the two independent extracts (*C. arabica* and *C. canephora*). A standard gel was generated out of the image with the highest spot number. Spot quantities of all gels were normalized by removing nonexpression-related variations in spot intensity; for that, the raw quantity of each spot in a gel was divided by the total quantity of all the spots in that gel and included in the standard. The final synthetic image was a Gaussian scan image that contained all the Gaussian spots with a defined volume and quality. All subsequent spot matching and analysis steps in PDQuest software were performed on Gaussian spots. The results were evaluated in terms of spot OD. Statistical analysis (Student's *t*-test) via PDQuest allowed the study of proteins that were significantly increased or decreased in the two sets of samples.

2.5 In situ digestion and extraction of peptides

Spots were carefully cut out from 2-D Coomassie stained gels using a razor blade and subjected to in-gel trypsin digestion according to the method of Shevchenko *et al.* [16] with minor modifications. The gel pieces were swollen in a digestion buffer containing 50 mM NH_4HCO_3 and 12.5 ng/ μL trypsin (modified porcine trypsin, sequencing grade; Promega, Madison, WI, USA) in an ice bath. After 30 min the supernatant was removed and discarded and 20 μL of 50 mM NH_4HCO_3 was added to the gel pieces. Digestion was allowed to proceed at 37°C overnight. The supernatant containing tryptic peptides was dried by vacuum centrifugation. Prior to mass spectrometric analysis, the peptide mixtures were redissolved in 10 μL of 5% formic acid (FA).

2.6 MALDI-TOF MS

Peptides were desalted and concentrated according to the method of Gobom *et al.* [17]. Home-made 5 mm nanocolumns were packed with POROS R2 resin (PerSeptive Biosystems, Framingham, MA, USA) in a constricted GELoader tip (Eppendorf, Hamburg, Germany). A syringe

was used to force liquid through the columns by applying gentle air pressure. The columns were equilibrated with 20 μL of 5% FA and the analyte solutions were added. The columns were washed with 20 μL of 5% FA and the bound peptides subsequently eluted directly onto the MALDI target with 0.5 μL CHCA solution (20 $\mu\text{g}/\text{mL}$ in ACN/0.1% TFA, 70:30 vol/vol). MALDI-MS was performed using a Biflex III instrument (Bruker-Daltonics, Bremen, Germany), equipped with delayed extraction. Positively charged ions were analyzed in reflectron mode, using delayed extraction. Typically, 200 shots were averaged to improve the S/N ratio. Spectra were calibrated using trypsin autolysis products (m/z 842.51 and 2211.10) as internal standards. In a few cases this was not possible and external calibration was carried out by using peaks from a peptide calibration standard II (Bruker-Daltonics). Protein identification was performed by searching in a nonredundant protein sequence database (NCBI nr) via the MASCOT program (<http://www.matrixscience.com>). The following parameters were used for database searches: Monoisotopic mass accuracy 100 ppm, missed cleavages 2, complete carbamidomethylation of cysteines and partial oxidation of methionines. For positive identification, the score of the result of $(-10 \times \text{Log}(P))$ had to be over the significance threshold level ($p < 0.05$).

2.7 RP-HPLC-ESI-MS/MS

Peptide mixtures were separated using an LC packings system autosampler/nanoHPLC (Amsterdam, The Netherlands). A sample volume of 8 μL was loaded by the autosampler and concentrated on a commercial trap column (Zorbax 300 SB-C18, 5 μm , 300 μm I.D. $\times$ 5 mm; Agilent Technologies, Palo Alto, CA, USA) at a flow rate of 10 $\mu\text{L}/\text{min}$. Separation was performed using a homemade capillary column (Zorbax 300 SB, C18, 3.5 μm , 75 μm id $\times$ 10 cm) and a flow rate of 200 nL/min. HPLC solvents contained 0.1% FA and either 2% ACN (solvent A) or 98% ACN (solvent B). The column was pre-equilibrated with 100% solvent A. Elution was performed by a multisegment gradient with a first step from 0 to 5% solvent B in 5 min, a second one from 5 to 12% solvent B in 2 min, a third one from 12% to 50% solvent B in 30 min and a last segment from 50 to 100% solvent B in 8 min. Peptides were eluted directly into a Bruker ion trap Esquire 3000+ (Bruker-Daltonics). The capillary voltage was 1.5–2 kV and a dry gas flow rate of 3 L/min was used with a temperature of 230°C. The scan range used was from 300 to 1800 m/z . Typically, data were collected using an automatic data-dependent scanning mode (isolation width 4 m/z ; fragmentation amplitude 1.15 V). The spectrometer sequentially conducts MS/MS on the two most abundant precursor ions detected in the full scan. For the MS/MS spectra eight averages were allowed. For protein identification the following parameters were adopted: Complete carbamidomethylation of cysteines and partial oxidation of methionines, peptide mass tolerance ± 1.2 Da, fragment mass tolerance ± 0.9 Da, missed cleavages 2.

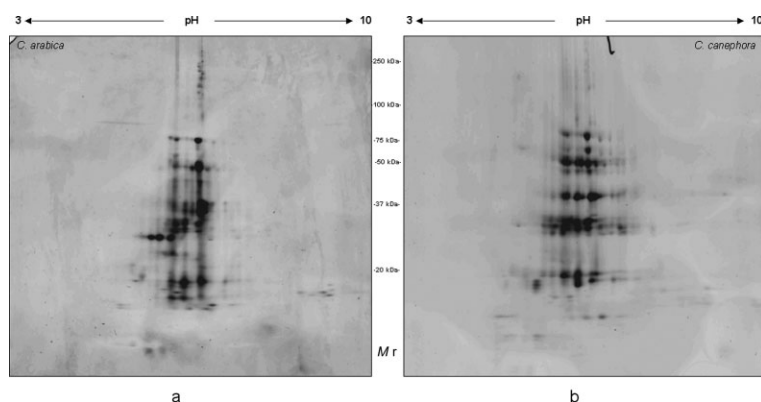


Figure 1. Representative maps of (a) *C. arabica* (var. Colombia) and (b) *C. canephora* (var. Indiano Robusta). First dimension: IPG pH 3–10 linear gradient. Second dimension: SDS-PAGE in a 7%T to 20%T porosity gradient.

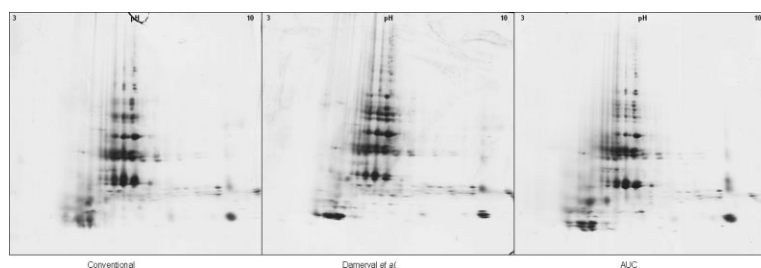


Figure 2. Examples of three 2-D maps of *C. canephora* (var. Indiano Robusta) obtained by extracting coffee grains with three different methods (from left to right: conventional method, method from Damerval *et al.* [11] and the AUC variant). The gels were stained with Sypro Ruby.

3 Results

Figure 1 gives representative maps of *C. arabica* (var. Colombia) (Fig. 1a) and *C. canephora* (var. Indiano Robusta) (Fig. 1b). The strings of spots characterizing the α and β arms, at 33 kDa and 22 kDa, respectively, are clearly visible, as well as their precursor at 55 kDa. The string of spots around 80 kDa could be oligomers (tetramers) of the β arm, as suggested by Rogers *et al.* [6]. Figure 2 shows three 2-D maps obtained by extracting the grains with the three different methods reported in section 2 (conventional method, method from Damerval *et al.* [11] and the AUC variant). In our hands, there does not seem to be any difference among the three protocols, therefore we adopted the well-engrained Damerval extraction procedure for all subsequent work. Figure 3 gives the results of the differential analysis obtained with PQQuest software. The 16 polypeptide chains present only in *C. canephora* (var. Indiano Robusta) are shown in A; panel B displays the five spots present only in *C. arabica* (var. Colombia); panel C shows eight polypeptide chains up-regulated (by factors of two to five) in *C. canephora* (var. Indiano Robusta) vs. *C. arabica* (var. Colombia) and panel D shows

the only protein we could detect up-regulated in *C. arabica* (var. Colombia). The pie chart in panel E summarizes the data on the differentially expressed proteins in the two green coffee species. A number of these proteins were identified by MALDI-TOF (Fig. 4) and ESI-MS when required. Their names and functions are given in Table 1.

4 Discussion

To our knowledge, this is the first attempt at a phylogenetic study of different coffee species *via* a differential proteomic approach. Given the number of proteins spots which were found to characterize, *via* their variation, the two species (a total of 30 proteins differentially expressed), it is hoped that, as more species are analyzed (even of closely-related lineages) it will be possible to differentiate them *via* this type of analysis. It should be remembered that this is the minimum number of spots that we found to be varied in the two species, since we adopted a medium-sensitivity stain (colloidal Coomassie Blue), which has about one order of magnitude less sensitivity than classical Silver-staining protocols [18]. We did not use Silver-staining since our aim was to detect spots which would be amenable to subsequent analysis by MS. It is well known that Silver-stains, due to the presence of formaldehyde or glutaraldehyde, are incompatible with MS analysis. Although recently aldehyde-free Silver-stains have been described [19–21], their sensitivity is much reduced, so that they do not seem to perform any better than the micellar Coomassie adopted here. Nevertheless, for identification purposes, when closely-related species are analyzed in the future, one could adopt well-controlled Silver-staining protocols or fluorescent stains to reveal many more polypeptide chains than presently reported and thus enhance the chances of detecting only subtle differences.

Our approach was also used by Rogers *et al.* [6]; however, their 2-D maps were only shown to highlight some differences between *C. arabica* and *C. canephora*, but no differential analysis with dedicated software was attempted. Interestingly, a completely different approach was attempted by Procida *et al.* [22] for obtaining fingerprints of native proteins from *C. arabica* and *C. canephora*, for differentiation of these two species as well as for characterizing differences among samples of the same variety from various plantations. They extensively dialyzed the green beans, ground them to a fine powder, mixed the powder with the MS matrix (sinapinic acid at saturation in ACN/water (50:50) in the presence of 0.1% TFA) and analyzed the entire extract by MALDI-TOF MS. In the analysis of 22 plantations of *C. arabica*, they constructed a matrix of 35 signals (peaks ranging in m/z from

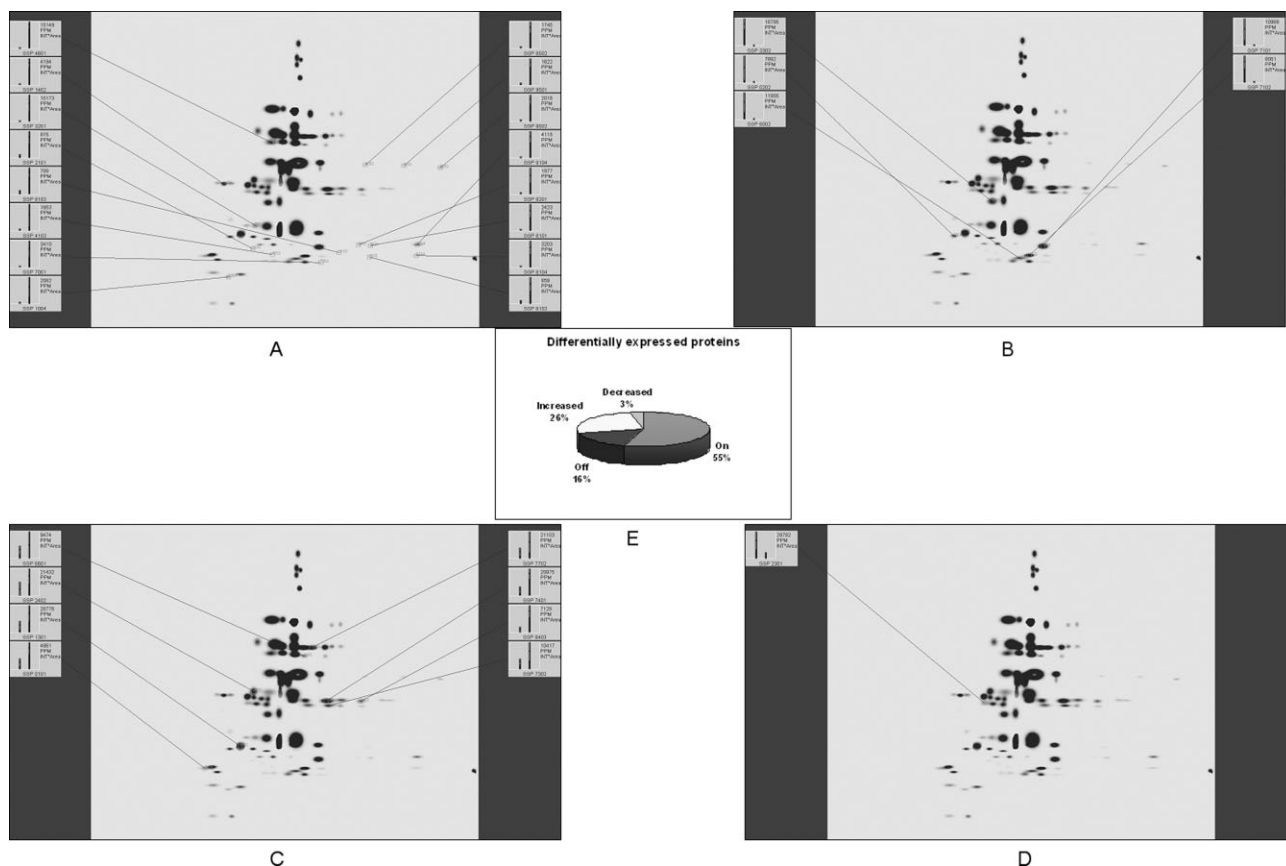


Figure 3. Variation in protein profiles between *C. arabica* (var. Colombia) and *C. canephora* (var. Indiano Robusta). (A) Visualization of the 16 polypeptide chains present only in *C. canephora* (var. Indiano Robusta). (B) Display of the five spots present only in *C. arabica* (var. Colombia). (C) Visualization of the eight polypeptide chains up-regulated (by factors of two to five) in *C. canephora* (var. Indiano Robusta) vs. *C. arabica* (var. Colombia) (increased). (D) Display of the only detectable protein up-regulated in *C. arabica* (var. Colombia) (decreased). (E) Summary of the data on the differentially expressed proteins in the green coffee species of *C. arabica* (var. Colombia) and *C. canephora* (var. Indiano Robusta).

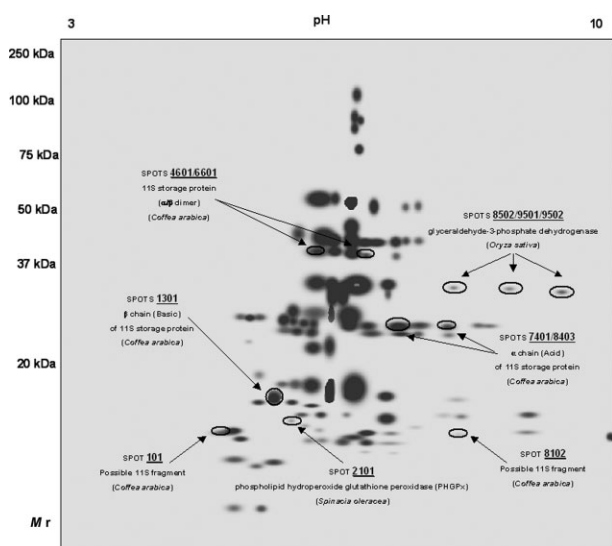


Figure 4. Master map of green coffee beans. The circled spots refer to 11 spots which could be identified by MS upon excision.

2438 up to 13 162) of which about a dozen appear to be present (with different intensities) in almost all of the 22 specimens. The other signals were characteristic of only a few cultivars. In the analysis of 17 plantations of *C. canephora*, they built a matrix of 30 signals (peaks ranging in m/z from 2515 up to 13 139), of which only five appear to be present (here too with different intensities) in almost all of the 17 specimens, the others being characteristic of just one or a few cultivars. Hierarchical classifications of *C. arabica* and *C. canephora* were obtained from these matrices by using the Ward's clustering method and presented in the forms of dendrograms. Even closely related plantations could be identified. This approach is certainly interesting and has the advantage of requiring a minimum of labor, since it exploits direct analysis of the entire powdered grains. However, it is impossible for us to compare the two sets of data for several reasons. First of all, Procida *et al.* [22] might also be looking at rather small peptides, which are in general not visible in 2-D maps, unless one uses special buffers in the SDS-PAGE dimension, as performed by Ludwig *et al.* [7]. Secondly, we

Table 1. Identified proteins from green coffee beans.

SSP number	Identified peptides		Theoretical M_r Calculated pI	NCBI nr data- base acces- sion number % Sequence coverage	Protein name and function	Variation (Indiano Robusta vs. Colombia)
	ESI-MS/MS ions	MALDI ions [M+H ⁺]				
8502	400.75 (+2) VVISAPSK	Not Analyzed	42 052.2 Da 6.41	gi 29150193 (<i>Oryza sativa</i>) 23%	Putative glyceraldehyde- 3-phosphate dehydroge- nase (GAPD is a key enzyme in intermediary metabolism catalyzing the oxidation and sub- sequent phosphorylation of substrate aldehydes to acyl phosphate)	ON
9501	404.25 (+2) TLLFGEK					
9502	406.27 (+2) VLPALNGK					
	464.8 (+2) KVISAPSK					
	653.25 (+2) DAPMFVVGVEK					
	661.31 (+3) DAPMFVVGVEK + Oxidation (M)					
	716.47 (+3) PDIDIVSNASCTTN					
	CLAPLAK					
	717.89 (+2) AASFNIIPSSTGAAK					
	749.95 (+2) VPTVDVSVVDLTVR					
2101	487.21 (+2) SVHEFVVR	Not Analyzed	19 047.6 Da 5.92	gi 20138151 (<i>Spinacia oleracea</i>) 22%	Probable phospholipid hy- droperoxide glutathione peroxidase (Protects cells and enzymes from oxidative damage, by catalyzing the reduction of hydrogen peroxide, lipid peroxides and or- ganic hydroperoxide, by glutathione)	ON
	491.74 (+2) AEYPIFDK					
	562.48 (+2) GNDVDLSIYK					
	674.34 (+2) VDVNGSNAPIYK					
7401	Not Analyzed	1054.53 LPHYSNVPK 1061.55 LNAQEPSFR 1320.71 FFLAGNPQQGGGK 1876.94 TQCNIQKLNAQEPSFR 3299.64 LQLFLPEYSE QEQQPQQQQGQQQQGVGR	30 428.7 Da 6.93 ^{a)}	gi 2979526 (<i>Coffea arabica</i>) 24% ^{c)}	11S storage protein Acid chain (α) (This is a seed storage protein. He- xamer; each subunit is composed of an acidic and a basic chain derived from a single precursor and linked by a disulfide bond. Belongs to the 11S seed storage protein (globulins) family)	Up-regulated 3.7
8403	527.77 (+2) LPHYSNVPK 531.11 (+2) LNAQEPSFR 660.82 (+2) FFLAGNPQQGGGK 1121.53 (+2) NIFSGFDDQLLA DAFNVDLK 1250.74 (+3) LQLFLPEY- SEQVQQPQQQQE QQQHGVGRGWR	927.37 FPDRHQQ 1054.48 LPHYSNVPK 1061.75 LNAQEPSFR 1320.68 FFLAGNPQQGGGK 2242.20 NIFSGFDDQLLA DAFNVDLK	31 418.8 Da 7.03 ^{a)}	gi 4127631 (<i>Coffea arabica</i>) 31% ^{d)}	11S storage protein Acid chain (α)	Up-regulated 6.2
8102	483.60 (+3) KFLAGNPQQGGGK 660.84 (+2) FFLAGNPQQGGGK 690.92 (+2) GLRLPHYSNVPK 1528.40.60 (+3) GDVLILLPGFTQW TYNDGDV PLTVALLDVA NEANQLDLQSR	Not Analyzed	51 195.3 Da 6.65 ^{a)}	gi 2979526 (<i>Coffea arabica</i>) 15%	Possibile 11S fragment	ON
101	618.88 (+2) GNARIQVVDHK 666.32 (+2) AGNEGFYVAFK 707. 63 (+2) TNDNAMINPLVGR 715.85 (+2) TNDNAMINPLVGR + Oxidation (M)	Not Analyzed	51 195.3 Da 6.65 ^{a, b)}	gi 2979526 (<i>Coffea arabica</i>) 8%	Possibile 11S fragment	Up-regulated 3.0

Table 1. Continued

SSP number	Identified peptides		Theoretical M_r Calculated pI	NCBI nr data- base access- ion number % Sequence coverage	Protein name and function	Variation (Indiano Robusta vs. Colombia)
	ESI-MS/MS ions	MALDI ions [M+H ⁺]				
1301	Not Analyzed	1331.84 AGNEGFEYVAFK 1427.05 IPILSSLQLSAER 1459.97 KAGNEGFEYVAFK 1584.04 SSFQISSEAEELK 1717.29 QGQLIIVPQYFAVIK 1899.24 LSENIGLPQEADVFNPR 1960.21 SSFQISSEAEELKYGR	20 784.5 Da 6.3 ^{a, b)}	gi 2979526 (<i>Coffea arabica</i>) 40% ^{e)}	11S storage protein Basic chain (β)	Up-regulated 2.3
4601	531.12 (+2) LNAQEPSFR 660.59 (+2) FFLAGNPQQGGGK 666.37 (+2) AGNEGFEYVAFK 715.85 (+2) TNDNAMINPLVGR 792.72 (+2) SSFQISSEAEELK	Not Analyzed	51 195.3 Da 6.65 ^{a)}	gi 2979526 (<i>Coffea arabica</i>) 13%	11S storage protein (α/β dimer)	ON
6601	Not Analyzed	1899.33 LSENIGLPQEADVFNPR 1427.08 IPILSSLQLSAER 717.32 QGQLIIVPQYFAVIK 584.03 SSFQISSEAEELK 061.47 LNAQEPSFR 380.56 LQLFLPEYS EQEQQQQQQQEQQQHGVGR	964.1 Da 6.73 ^{a)}	gi 3641256 (<i>Coffea arabica</i>) 23%	11S storage protein (α/β dimer)	Up-regulated 2.4

a) Calculated pI value according to [6]

b) For possible 11S fragments we reported the calculated M_r and pI of the entire legumin

c) Calculated considering the amino acid sequence corresponding to the acid chain 28–300 aa of the entire precursor (gi|2979526)

d) Calculated considering the amino acid sequence corresponding to the acid chain 27–306 aa of the entire precursor (gi|4127631)

e) Calculated considering the amino acid sequence corresponding to the basic chain 301–487 aa of the entire precursor (gi|2979526)

The comparison is performed between the matched spots in *C. canephora* (var. Indiano Robusta) and *C. arabica* (var. Colombia). A threshold value of 2 (which corresponds to a variation of 100%) was chosen as a meaningful variation in the comparison of Indiano proteins vs. Colombia proteins. The numbers in the last column refer to folds of change.

might be seeing a set of quite different polypeptide chains, due to the fact that our samples are solubilized in strong cocktails containing the classical mixture of urea/thiourea and surfactants, whereas their solubilization mixture is just the solvent for sinapinic acid, which may not bring into solution whole classes of proteins. Thirdly, whereas our proteins, as extracted from a 2-D map, are characterized by a pI, an intact mass in MALDI-TOF (as well as peptide masses for identification purposes) and, when required, a lead sequence *via* ESI-MS, their peaks are only characterized by an m/z signal, which unfortunately cannot tell much about the properties of the polypeptide chain accompanying the signal.

5 Concluding remarks

Although preliminary in nature, these studies will be quite valuable in view of extending them to quality control: *e.g.* fraud detection (such as mixing of beans of different origins and of different commercial value), food adulteration (for example, if coffee beans are mixed with husks, hulls, or parchments (non-bean parts of the coffee cherry) the pro-

cessed soluble coffee will contain high levels of the sugar xylose), and food products containing material derived from a genetically modified organism (GMO).

Scientists in the food and beverage industry are faced with many different quality control tasks, such as making sure that flavors meet certain standards, identifying changes in process parameters that may lead to a change in quality, detecting adulterations in ingredients and identifying the geographical origin of raw materials. Food scientists who work for regulatory agencies, such as the Food and Drug Administration, are interested in detecting economic fraud due to product substitution and adulteration, as well as health risks from possible contamination. In this scenario, the proteomic approach could be a valuable method to assist in quality control.

P. G. R. is supported by grants from MIUR (COFIN 2003 and FIRB2001, no. RBNE01KJHT). M. T. G. A. would like to thank the Secretaría de Estado de Educación y Universidades and the European Social Fund for her post-doctoral fellowship. L. Z. is supported by Progetto Strategico 1999 (Area 2.2).

6 References

- [1] Capodici, S., Invernizzi, C., *Conoscere il Caffè*, Eusebianum, Milano, Italy 1983, pp. 27–30.
- [2] Charrier, A., Eskes, A. B., in: Charrier, A., Jacquot, M., Hamon, S., Nicolas, D. (Eds.), *L'Amélioration des plantes Tropicales*, Cirad Orstrom, Toulouse, France 1997, pp. 171–196.
- [3] Flament, I., *Coffee Flavour Chemistry*, Wiley, New York, USA 2002, pp. 79–99.
- [4] Brinkman, U. A., Vreuls, R. J. J., *J. Chromatogr. A* 2003, **1019**, 1–285.
- [5] Holscher, W., Steinhart, H., in: Parliament, T., Morello, M., McGorin, R. (Eds.), *Thermally Generated Flavours*, ACS Symposium Series 543, American Chemical Society, Washington, USA 1992, pp. 207–217.
- [6] Rogers, W. J., Bezard, G., Deshayes, A., Meyer, I. *et al.*, *Plant. Physiol. Biochem.* 1999, **37**, 261–272.
- [7] Ludwig, E., Lipke, U., Raczek, U., Jäger, A., *Eur. Food Res. Technol.* 2000, **211**, 111–116.
- [8] Schägger, H., von Jagow, G., *Anal. Biochem.* 1987, **166**, 368–379.
- [9] Montavon, P., Duruz, E., Rumo, G., Pratz, G., *J. Agric. Food Chem.* 2003, **51**, 2328–2334.
- [10] Montavon, P., Mauron, A. F., Duruz, E., *J. Agric. Food Chem.* 2003, **51**, 2335–2343.
- [11] Damerval, C., De Vienne, D., Zivy, M., Thielllement, H., *Electrophoresis* 1986, **7**, 52–54.
- [12] Herbert, B., Galvani, M., Hamdan, M., Olivieri, E. *et al.*, *Electrophoresis* 2001, **22**, 2046–2057.
- [13] Righetti, P. G., *Immobilized pH Gradients: Theory and Methodology*, Elsevier, Amsterdam, The Netherlands 1990, pp. 206–220.
- [14] Steinberg, T. H., Chernokalskaya, E., Berggren, K., Lopez, M. F. *et al.*, *Electrophoresis* 2000, **21**, 486–496.
- [15] Garrells, J. I., Farrar, J. T., Burwell, C. B., in: Cedlis, J. E., Bravo, R. (Eds.), *Two-Dimensional Gel Electrophoresis of Proteins*, Academic Press, Orlando, USA 1984, pp. 37–91.
- [16] Shevchenko, A., Wilm, M., Vorm, O., Mann, M., *Anal. Chem.* 1996, **68**, 850–858.
- [17] Gobom, J., Nordhoff, E., Mirgorodskaya, E., Ekman, R., Roepstorff, P., *J. Mass Spectrom.* 1999, **34**, 105–116.
- [18] Merrill, C. R., Washart, K. M., in: Hames, B. D., (Ed.), *Gel Electrophoresis of Proteins: a Practical Approach*, Oxford University Press, Oxford, England 1998, pp. 53–91.
- [19] Mortz, E., Krogh, T. N., Vorum, H., Goerg, A., *Proteomics* 2001, **1**, 1359–1363.
- [20] Sinha, P., Poland, J., Schnoelzer, M., Rabilloud, T., *Proteomics* 2001, **1**, 835–840.
- [21] Yan, J. X., Harry, R. A., Spibey, C., Dunn, M. J., *Electrophoresis* 2000, **21**, 3657–3665.
- [22] Procida, G., Campisi, B., Seraglia, R., Traldi, P., *Rapid Commun. Mass Spectrom.* 2003, **17**, 140–148.

PARTE IV

**“Analisi delle proteine S-nitrosilate
espresse in *Arabidopsis thaliana* nella
risposta all’infezione del patogeno
Pseudomonas syringae pv. *tomato*”**

ABSTRACT

L'ossido nitrico (NO) è stato scoperto essere nelle piante una molecola ubiquitaria che agisce nella regolazione dello sviluppo e come segnale chiave nella risposta di resistenza ai patogeni. Sebbene l'NO è implicato in un gran numero di processi fisiologici, ancora poco si conosce riguardo i suoi target molecolari. In recenti studi è stato dimostrato che l'NO, nei tessuti animali, può S-nitrosilare i residui cisteinici in proteine target alterando, potenzialmente, la loro funzionalità. Questo meccanismo specifico, regolato ed ubiquitario può giocare un importante ruolo regolatorio alla pari di altre modifiche post-traduzionali come la fosforilazione. E' dimostrato che un numero sempre più crescente di proteine animali sono soggette ad S-nitrosilazione *in vivo*, tra queste l'emoglobina, la creatina chinasi, la gliceraldeide-3-fosfato e diverse caspasi. Un recente studio su *Arabidopsis thaliana* ha messo in evidenza che numerose proteine possono essere S-nitrosilate anche nelle piante.

In questo lavoro si è utilizzato un approccio di tipo proteomico per analizzare le proteine differenzialmente S-nitrosilate durante l'interazione incompatibile pianta-patogeno. Foglie di *A. thaliana* sono state infiltrate con *Pseudomonas syringae* pv. *tomato* recante il gene avirulento *AvrB*. Al fine di rilevare le proteine S-nitrosilate è stato utilizzato il metodo "biotin-switch", con il quale le cisteine nitrosilate vengono biotinilate dopo aver preventivamente bloccato i gruppi tiolo liberi con un agente metiltiolante. Le proteine biotinilate sono state purificate su NeutrAvidina agarosio e separate attraverso elettroforesi bidimensionale (2-DE). Gli spot ottenuti, rivelati con il colorante fluorescente Sypro Ruby, dopo digestione triptica sono stati analizzati mediante l'uso di uno spettrometro di massa MALDI-TOF/TOF. L'identificazione dell'intero complemento proteico S-nitrosilato e le conseguenze funzionali di tale modificazione, sono essenziali per la comprensione del meccanismo di azione dell'NO e dei percorsi di segnalazione che scaturiscono dal suo rilascio.

CAPITOLO 1

INTRODUZIONE

1.1 L'ossido nitrico (NO)

L'ossido nitrico (NO) è una molecola lipofila altamente reattiva, con una chimica molto complessa che coinvolge la formazione di svariate forme ossido-riduttive con diversa reattività [1]. L'NO si presenta infatti in tre forme soggette ad interconversione in seguito al guadagno o alla perdita di un elettrone: l'NO radicalico ($\text{NO}^\bullet$), il catione nitrosonio (NO^+) e l'anione nitrosile (NO^-) (Fig. 1.1).

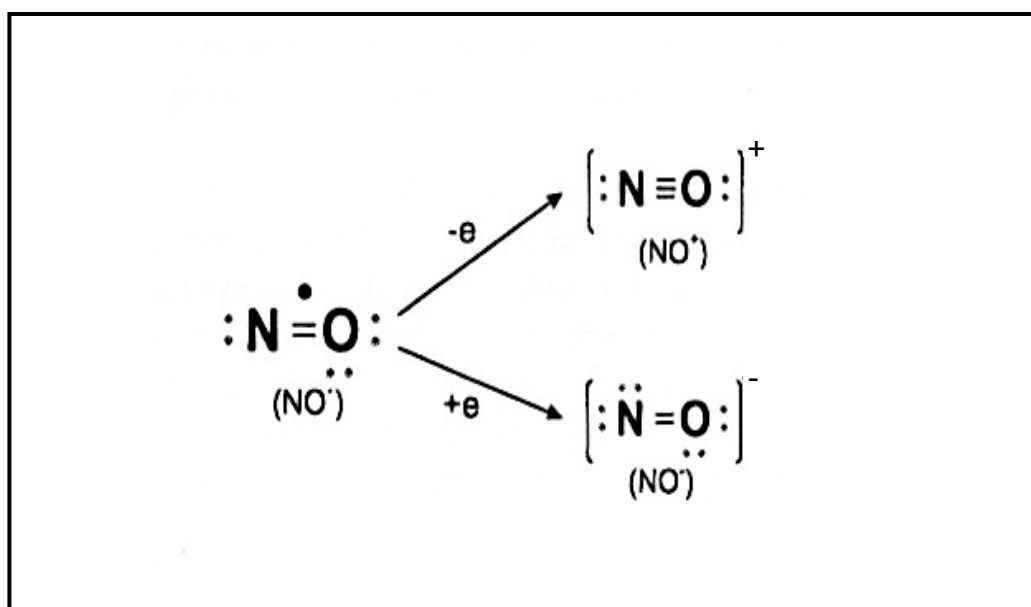


Figura 1.1 Rappresentazione delle possibili forme di conversione ossido-riduttiva dell'NO [2]. L'NO radicalico ($\text{NO}^\bullet$) è rapidamente ossidato dalla rimozione di un elettrone formando il catione nitrosonio (NO^+), oppure è ridotto dall'aggiunta di un elettrone formando l'anione nitrosile (NO^-).

La sua funzione ubiquitaria come molecola bioattiva riscontrata in procarioti, animali e piante, è fondamentale nello sviluppo dell'organismo durante l'intero ciclo di vita [3, 4].

A causa della sua alta reattività, molte delle reazioni che lo vedono implicato dipendono dalla sua concentrazione, dallo stato ossidativo cellulare e dalla disponibilità delle molecole bersaglio, piuttosto che da una specifica catalisi

enzimatica [5]. Queste reazioni regolano la sua disponibilità e la sua funzione di segnale, ma determinano anche la sua tossicità. L'NO può reagire rapidamente con l'ossigeno (O_2) formando diossido di azoto (NO_2) o con il radicale libero superossido ($O_2^{\cdot-}$), formando perossinitrito ($ONOO^{\cdot-}$) il quale direttamente, o tramite la formazione del radicale idrossile ($\cdot OH$), può danneggiare molecole biologicamente importanti come DNA, membrane e proteine [6]. L'NO, può esercitare un'azione protettiva nei confronti di stress ossidativi [7] basata sull'abilità regolativa del livello e della tossicità delle specie reattive dell'ossigeno [2] e sulla regolazione dell'omeostasi ossido-riduttiva cellulare [8].

Le funzioni dell'NO sono molteplici e dipendono dall'organismo considerato e dal tipo di cellula in cui viene prodotto. In animali è stato dimostrato il suo coinvolgimento in numerosi processi fisiologici quali la regolazione della risposta immunitaria ed infiammatoria, il mantenimento della pressione sanguigna nel sistema cardio-vascolare, la trasmissione neuronale e l'apoptosi [9-11]. Nei sistemi vegetali la produzione di NO accompagna l'intero ciclo di vita della pianta, dalla germinazione del seme [12] alla crescita del tubetto pollinico durante la fecondazione [13]. Numerose e varie sono le funzioni che esso svolge sia come molecola citotossica sia come molecola segnale in grado anche di attivare protezione cellulare [14].

Molte delle azioni biologiche attivate da NO dipendono dalla sua capacità di reagire con metalli di transizione o di modificare residui aminoacidici delle proteine. Alcune di queste modificazioni (nitrazione della tirosina, ossidazione della metionina) sono irreversibili e potenzialmente dannose; altre invece, come la S-nitrosilazione di residui cisteinici o la nitrosilazione del gruppo eme, sono reversibili e possono regolare finemente l'attività della proteina [15]. Notevole importanza è attribuita alla capacità dell'NO di nitrosilare il gruppo eme di svariate proteine. L'esempio meglio caratterizzato è la nitrosilazione del gruppo eme della guanilato ciclastasi animale che comporta l'attivazione enzimatica con conseguente aumento della quantità di GMP ciclico (cGMP), una molecola coinvolta nella regolazione di numerose cascate di trasduzione del segnale [16]. La regolazione dei livelli di cGMP è responsabile di molte delle risposte NO-dipendenti, tuttavia recenti studi hanno focalizzato l'attenzione su una nuova modificazione post-traduzionale, indotta da NO, che svolge un ruolo determinante nel modulare la sua funzione di segnale: la nitrosilazione dei residui cisteinici o S-nitrosilazione.

1.2 La S-nitrosilazione

La S-nitrosilazione o S-nitrosazione [17] è una modificazione post-traduzionale in cui una molecola di NO si lega al gruppo tiolico di una cisteina [18, 8] con formazione di un nitrosotio, o tionitrito estere ($R-S-N=O$; Fig. 1.2).

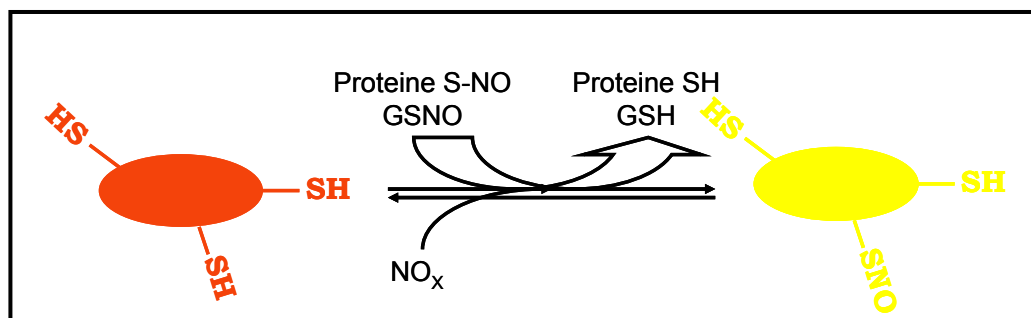


Figura 1.2 Rappresentazione schematica della S-nitrosilazione di una proteina.

La chimica che regola la S-nitrosilazione è ancora poco caratterizzata, nonostante sia ormai evidente che l'NO da solo è in grado di nitrosilare i metalli di transizione, sembra che siano principalmente le specie reattive derivanti dall'NO a S-nitrosilare le cisteine [17].

La trans-nitrosilazione, ovvero il passaggio del gruppo NO a tioli liberi proteici da parte di complessi metallo-NO o di altri S-nitrosotio, come l'S-nitrosoglutatione (GSNO), cisteine libere o altre proteine S-nitrosilate, richiede il contatto diretto donatore-accettore [19, 20]. Dati presenti in letteratura indicano inoltre l'esistenza di un equilibrio cellulare, basato sul trasferimento di gruppi NO, tra tioli a basso peso molecolare ed i tioli delle proteine [21, 22]. Ad oggi non sono ancora stati descritti specifici meccanismi enzimatici che siano in grado di catalizzare la S-nitrosilazione o la denitrosilazione, nonostante numerosi enzimi possano promuovere l'una o l'altra. Un esempio è rappresentato dalla superossido dismutasi Cu, Zn dipendente (Cu, Zn-SOD), che catalizza la S-nitrosilazione dell'emoglobina ad opera dell'NO in soluzione o dell'NO derivante dal GSNO [23, 24]. Esistono anche numerose evidenze che dimostrano come le interazioni proteina-proteina con le NO transferasi, inducano diverse proteine ad essere soggette a trans-nitrosilazione [15]. L'esempio meglio descritto riguarda il recettore N-metil-D-aspartato (NMDAR) [25].

Caratteristica fondamentale della S-nitrosilazione è la specificità di substrato. Fattori che influiscono sulla specificità della S-nitrosilazione sono le

interazioni elettrostatiche che controllano la nucleofilicità del gruppo tiolico (pKa) [26], la compartimentalizzazione idrofobica [27], i regolatori allosterici che possono modulare l'accessibilità del tiolo o la sua reattività [28, 29] e le interazioni con molecole transnitrosilanti. Le interazioni elettrostatiche, ad esempio, che una proteina può instaurare con diversi tipi di ioni (Ca_2^+ , Mg_2^+ , H^+) e con l' O_2 sono importanti nel modulare la specificità di questa selezione promuovendo la S-nitrosilazione di una cisteina rispetto ad un'altra. Tali interazioni elettrostatiche possono inoltre causare modificazioni strutturali allosteriche che sono in grado di promuovere la S-nitrosilazione o la denitrosilazione [15]. La specificità della S-nitrosilazione viene modulata anche dalla nucleofilicità del gruppo tiolico; due esempi ben caratterizzati sono la presenza di un motivo acido-base adiacente alla cisteina (Fig. 1.3) e, in metallo proteine contenenti Zn_2^+ , la coordinazione dei tioli delle cisteine con lo ione metallico. Tali motivi possono portare ad una significativa diminuzione del pKa del gruppo tiolico, favorendo la S-nitrosilazione [30].

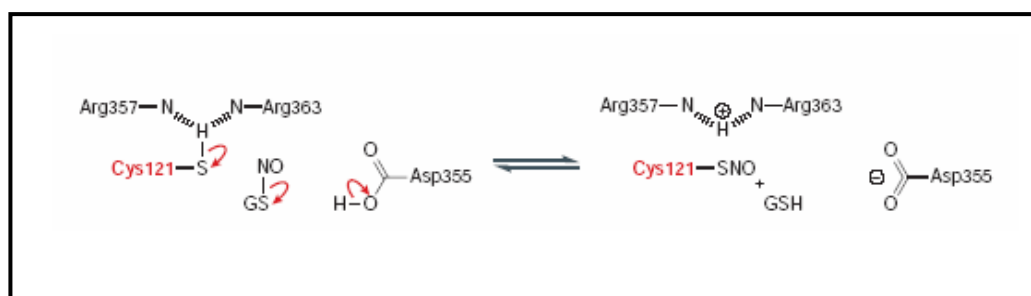


Figura 1.3 S-nitrosilazione indotta da GSNO della Cys 121 della metionina adenosiltransferasi (MAT) epatica, mediata dal dominio acido-base. Nella struttura tridimensionale il tiolo della Cys 121 è a breve distanza dai gruppi basici dell'Arg 357 e dell'Arg 363 e dal gruppo acido dell'Asp 355. La sostituzione di ciascuno di questi residui blocca la S-nitrosilazione della cisteina. Le interazioni elettrostatiche con i residui basici diminuiscono la pKa della Cys 121, mentre l'interazione del residuo di Asp 355 aumenta la pKa del GSNO e promuove la liberazione di NO^+ . La catalisi concertata acido-base promuove così la transnitrosilazione di MAT [31].

Un altro motivo proteico in grado di favorire la S-nitrosilazione è il dominio idrofobico. Regioni interne delle proteine altamente idrofobiche promuovono, infatti la S-nitrosilazione di residui cisteinici ivi presenti, ad opera degli ossidi dell'azoto (NO_2 , N_2O_3 , N_2O_4). Tale meccanismo rappresenta una forma di autocatalisi localizzata (Fig. 1.4) [32].

Nel corso degli ultimi dieci anni i substrati proteici della S-nitrosilazione identificati sono cresciuti ad oltre cento unità (una lista è disponibile sul sito: <http://www.cell.com/cgi/content/full/106/6/675/DC1>). I primi studi sono stati

realizzati in vitro inducendo la S-nitrosilazione tramite sorgenti esogene di NO (es: GSNO). Studi recenti hanno invece dimostrato la sua implicazione diretta in sistemi cellulari intatti. La S-nitrosilazione sembra svolgere un ruolo fondamentale nei meccanismi che modulano la trasduzione del segnale. Il suo controllo spaziotemporale, infatti, conferisce specificità agli effetti derivanti dall'NO e può mediare i segnali basati sull'alterazione dello stato ossido-riduttivo cellulare.

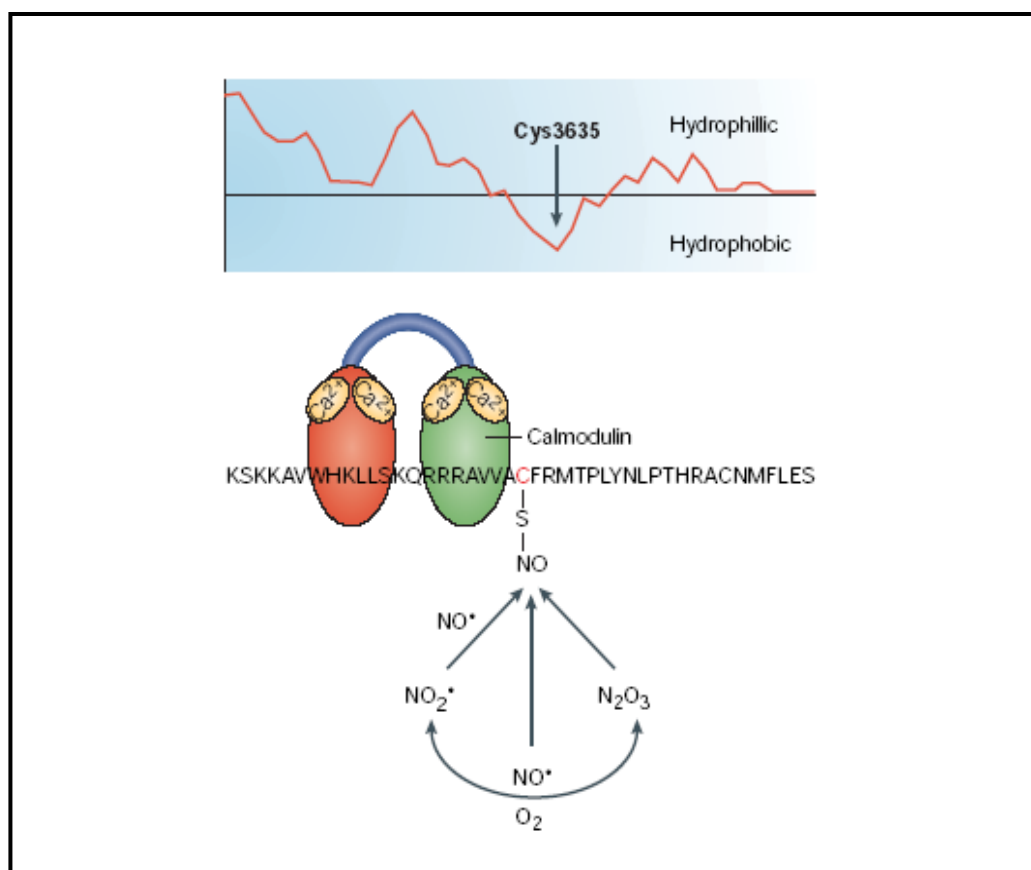


Figura 1.4 Rappresentazione della S-nitrosilazione della Cys 3635 del recettore rianodinico/canale di rilascio del calcio del muscolo scheletrico (RyR1). La Cys 3635 è intercalata all'interno di un dominio fortemente idrofobico (come evidenziato dal grafico) legante calmodulina. Tale dominio deriva dalla struttura tridimensionale della proteina e favorisce la S-nitrosilazione della cisteina ad opera degli ossidi idrofobici dell'azoto quali NO, N₂O₃ e NO₂ [4].

Il ruolo nella trasduzione del segnale da parte della S-nitrosilazione presenta molti aspetti comuni con un altro importante meccanismo di trasmissione del segnale, la fosforilazione. Entrambi i meccanismi comportano modificazioni covalenti, reversibili e specifiche delle proteine e sono in grado di mediare le risposte delle cellule ai segnali ambientali [33].

Gli effetti della S-nitrosilazione sono correlati alla modificazione di cisteine catalitiche o regolative che possono portare all'attivazione o all'inibizione dell'attività della proteina e, tramite la regolazione delle interazioni proteina-proteina, ad una specifica localizzazione subcellulare.

1.3 Metodologie di rivelazione per le proteine S-nitrosilate

L'analisi della S-nitrosilazione ha importanti implicazioni metodologiche in quanto il legame S-NO è influenzato dalla presenza di gruppi stabilizzanti o destabilizzanti, da sorgenti di protoni e elettroni e da altri aspetti della struttura proteica [15]. Esistono varie metodiche per misurare la quantità di proteine S-nitrosilate. Alcune si basano sulla misura diretta dei legami S-NO, altre invece su metodi indiretti che rilevano le specie che si formano dopo la rottura del legame S-NO.

Tecniche di misurazione diretta:

- a) misura del caratteristico picco di assorbanza a 340 nm;
- b) spettrometria di massa a ionizzazione elettrospray [34, 35];
- c) cromatografia liquida seguita da rilevazione elettrochimica (HPLC + ECD) [36];
- d) analisi NMR (^{15}N) [37].

Tecniche di misurazione indiretta:

- a) rilevazione tramite chemiluminescenza dell'NO rilasciato dagli S-nitrosotoli in seguito a:
 - taglio fotolitico [36];
 - riduzione utilizzando Cu^+ /cisteina a pH 6 [38];
 - riduzione utilizzando I_2/I^- [39];
- b) rilevazione del nitrito prodotto dal taglio del legame S-NO indotto da HgCl_2 tramite:
 - metodo colorimetrico, "Griess assay" [40];
 - metodo fluorimetrico, DAN [40, 41].

I metodi sopra elencati permettono di misurare l'ammontare totale di S-nitrosotoli nel campione. Una modifica della tecnica basata sulla rilevazione tramite ozono chemiluminescenza permette di rilevare le singole proteine S-nitrosilate nelle cellule o nei tessuti: la proteina di interesse viene immunoprecipitata con anticorpi specifici, ridotta con Cu^+ /cisteina o I_2/I^- e successivamente viene misurata la chemiluminescenza della soluzione [42, 43]. La sensibilità di questo approccio è

tuttavia limitata dall'ammontare di proteina immunoprecipitata e permette lo studio della specifica S-nitrosilazione di una sola proteina singolarmente. Recentemente per lo studio della S-nitrosilazione del fattore N-ethylmaleimide-sensitive [44] è stato utilizzato un anticorpo commerciale che permette l'immunorilevazione o l'immunoprecipitazione delle proteine contenenti cisteine S-nitrosilate (anti-CysNO, Calbiochem). Tuttavia, il metodo che finora ha permesso di ottenere i migliori risultati nello studio delle proteine S-nitrosilate è il "biotin switch" [45]. Tale metodo prevede la specifica conversione delle cisteine S-nitrosilate in cisteine biotinilate e la successiva immunorilevazione con anticorpi anti-biotina (Fig. 1.5).

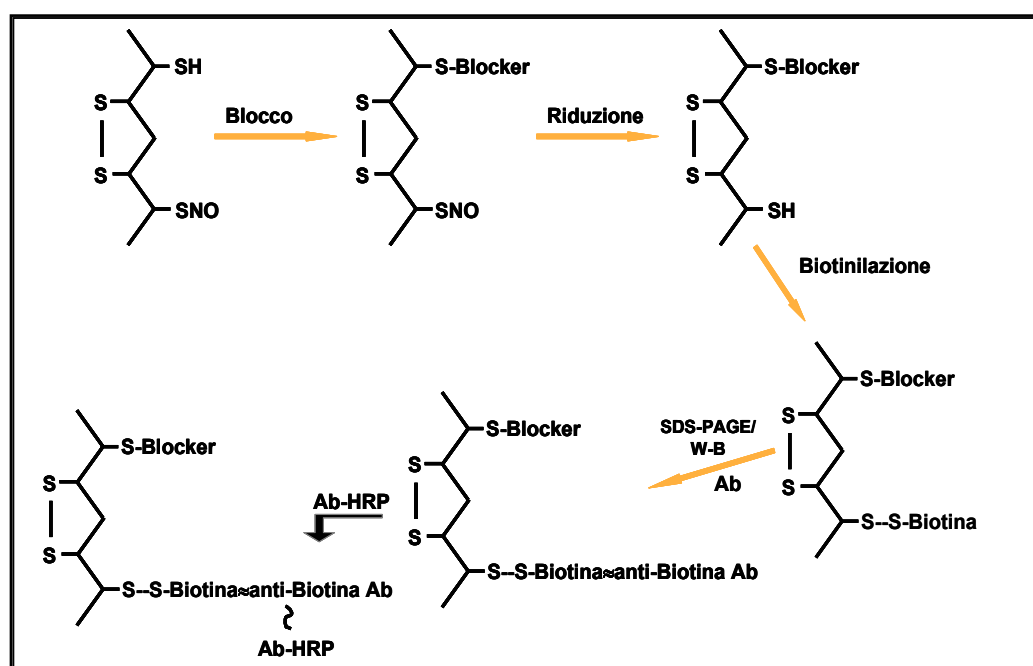


Figura 1.5 Rappresentazione schematica del metodo "biotin switch" di rilevazione delle proteine S-nitrosilate..

Una variante del metodo prevede l'immunoprecipitazione delle proteine biotinilate utilizzando streptavidina agarosio. Tali proteine sono quindi rese disponibili per analisi di tipo proteomico.

Il rapido sviluppo delle tecniche di analisi proteomica ha reso possibile lo studio del profilo proteico di cellule in diverse condizioni, e lo studio delle modificazioni post-traduzionali che possono alterare la struttura e la funzione delle proteine. L'analisi bidimensionale rappresenta un valido supporto allo studio del proteoma, consentendo una maggior separazione delle proteine e l'analisi di campioni biologici complessi. Il metodo più efficace è basato sul frazionamento

massa/carica, ottenuto combinando l'isoelettrofocalizzazione (IEF) nella prima dimensione ad una elettroforesi in presenza di sodio dodecil-solfato (SDS-PAGE) nella seconda dimensione. Tale tecnica raggiunge il teorico potere di risoluzione di 10.000 macchie proteiche individuali per mappa, anche se in pratica è possibile risolverne circa 3.000 [46]. Per quanto riguarda la S-nitrosilazione lo studio della porzione di proteoma S-nitrosilato (definito anche nitrosiloma) [47] è stato effettuato in estratti di cervello [45], in cellule mesangiali [47] e in cellule endoteliali [48]

1.4 Produzione di NO nelle piante

L'NO in pianta può essere prodotto sia non enzimaticamente sia tramite catalisi enzimatica. Quest'ultima è stata ben caratterizzata in sistemi animali in cui la biosintesi dell'NO avviene ad opera di ossido nitrico sintasi (NOS), che ossidano la L-arginina con produzione di L-citrullina e NO (Fig. 1.6). Si tratta di una reazione molto complessa che richiede NADPH e O_2 come co-substrati e flavinadeninadinucleotide (FAD), flavinammononucleotide (FMN) e tetraidrobiopterina (BH_4) come cofattori [49]. Sono state identificate tre isoforme di NOS animale: la iNOS (inducibile) è responsabile della produzione di NO durante la risposta immunitaria ed infiammatoria; la eNOS (endoteliale) e la nNOS (neuronale) sono invece espresse costitutivamente.

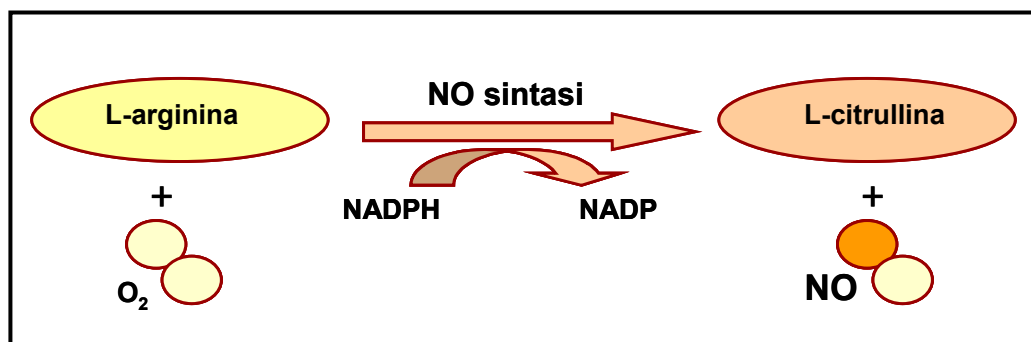


Figura 1.6 Produzione di NO da parte della NOS..

Nel corso degli ultimi anni, diversi studi hanno dimostrato l'esistenza di un'attività "NOS like" anche in sistemi vegetali. In molte specie vegetali e in diversi compartimenti subcellulari è stata osservata la reazione di ossidazione dell'arginina a citrullina e NO, analoga a quella catalizzata dalle NOS animali [50]. A supporto della tesi dell'esistenza di proteine NOS-like nelle piante è stata

dimostrata inoltre l'inibizione della produzione di NO in seguito al trattamento con inibitori della NOS animale [51], mentre analisi immunologiche e immunocitochimiche, utilizzando anticorpi anti-NOS di mammifero, hanno confermato tali evidenze sperimentali [52]. L'esistenza di una NOS vegetale è stata, tuttavia, a lungo controversa, dato che nel genoma di *Arabidopsis thaliana* non è stata trovata alcuna sequenza omologa alle NOS di mammifero [53]. È stato inoltre osservato che gli anticorpi anti-NOS utilizzati in pianta riconoscono aspecificamente numerose proteine che non presentano alcuna correlazione con la NOS [54].

Recentemente, in *A. thaliana* è stato identificato un gene espresso costitutivamente e codificante per una proteina con attività "NOS like" (AtNOS1). Tale gene non presenta omologie di sequenza con le NOS animali, ma è simile ad una proteina di *Helix pomatia* coinvolta nella sintesi dell'NO. L'attività di AtNOS1 dipende da NADPH, Cam, Ca_2^+ e, al contrario di quelle animali, è indipendente da BH_4 , FAD e FMN. Tale enzima svolge un ruolo vitale nella crescita della pianta, nella fertilità, nella regolazione dell'apertura/chiusura degli stomi, nel sistema di segnalazione regolato da ormoni e nella risposta di resistenza ai patogeni [55].

Un altro enzima che è stato dimostrato essere implicato nella sintesi dell'NO nelle piante è la nitrato riduttasi (NR) [56]. Essa svolge un ruolo fondamentale nell'assimilazione dell'azoto e sia *in vitro* che *in vivo* è stata dimostrata la sua capacità di trasferire un solo elettrone dal NADPH al nitrito per produrre NO [57] (Fig. 1.7).

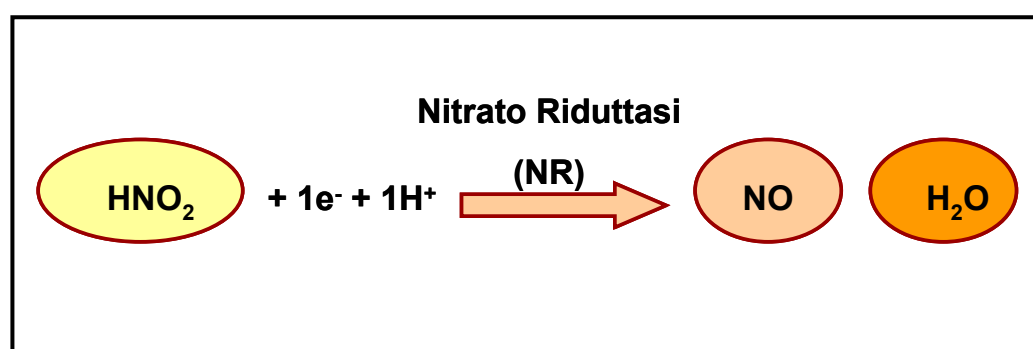


Figura 1.7 Produzione di NO da parte della NR.

Si ritiene che la fine regolazione trascrizionale e post-traduzionale dell'attività della NR [58], oltre che nella regolazione dell'assimilazione del nitrato, abbia un ruolo fondamentale nel controllo dell'accumulo di NO nei tessuti e, quindi, dello stress nitrosativo [59]. L'importanza di tale meccanismo di produzione

dell'NO è stata confermata da evidenze farmacologiche, fisiologiche e genetiche che indicano l'NO prodotto dalla NR come il responsabile della regolazione di numerosi meccanismi fisiologici [60]. Altre potenziali sorgenti enzimatiche di NO sono la xantina ossidasi e il citocromo P450 [61].

Per quanto riguarda la produzione non enzimatica di NO è stato dimostrato che una lenta e spontanea liberazione di NO avviene a pH neutro a partire da nitrito [59]. Un ambiente acido e la presenza di agenti riducenti, come ascorbato o composti fenolici (es: carotenoidi), molto abbondanti nelle cellule vegetali, sembrano accelerare la velocità di questa reazione [62]. Recentemente è stato inoltre osservato che la sintesi di NO nei semi avviene in modo non enzimatico attraverso la riduzione del nitrito presente nell'apoplasto [62].

1.5 Ruolo dell'ossido nitrico nei sistemi vegetali

Anche nei sistemi vegetali, come in quelli animali, l'NO svolge un importante ruolo come molecola segnale durante l'intero ciclo di vita della pianta. In particolare, esso svolge importanti funzioni durante lo sviluppo radicale, lo sviluppo dei tessuti vascolari, la fioritura, la fecondazione e la germinazione dei semi [1]. Auxine e NO, infatti, coniugando un meccanismo cGMP-dipendente e cGMP-indipendente, mediano la formazione di un nuovo sistema di radici quando la radice primaria è rimossa [63, 64]; inoltre, la localizzazione asimmetrica dell'auxina dovuta al gravitropismo, porta ad un accumulo asimmetrico di NO nelle radici orizzontali, che media la curvatura della radice in risposta alla gravità [65]. A livello di sviluppo dei tessuti vascolari l'NO regola la morte cellulare programmata e la lignificazione durante il differenziamento del tessuto xilematico. Recenti studi hanno poi dimostrato il coinvolgimento dell'NO nella repressione della fioritura autonoma e dipendente del fotoperiodo. L'NO agisce inoltre anche come modulatore della crescita del tubetto pollinico nel corso della fecondazione, regolandone l'orientamento all'interno del pistillo tramite un meccanismo cGMP-dipendente [13].

L'allungamento degli ipocotili e degli internodi [14], la germinazione dei semi [12], la chiusura degli stomi indotta da acido abscissico [60, 2] sono altri processi fisiologici dove è stato dimostrato un ruolo chiave dell'NO.

Esso riveste notevole importanza in molti tipi di risposta della pianta anche a stress biotici e abiotici. Per quanto riguarda gli stress abiotici l'NO è coinvolto nella resistenza a stress osmotico, a stress salino, ad alte temperature e ad

irradiazioni UV-B [14]. E' stato dimostrato, inoltre, che trattamenti con bassi livelli di NO aiutano la tolleranza alla disidratazione nei semi di grano [2], proteggono piante di pomodoro dai danni indotti da stress ossidativo [14] e inducono la resistenza a stress salini aumentando il rapporto K^+/Na^+ [66]. Per quel che riguarda, invece, la risposta a stress biotici, l'NO è coinvolto nella resistenza della pianta ai patogeni, in particolare durante le interazioni di tipo incompatibile [67, 68].

1.6 NO nella risposta di ipersensibilità ai patogeni

Le piante, durante l'intero ciclo di vita, entrano in contatto con una miriade di microorganismi. Solo una parte di questi, però, sono potenzialmente patogeni e raramente la loro interazione con la pianta degenera in malattia. Numerose e varie sono infatti le strategie di difesa che la pianta è in grado di adottare. Le barriere fisiche che impediscono l'entrata in contatto del patogeno con l'ospite, e l'accumulo di composti chimici ed enzimatici spesso liberati in seguito alla rottura del vacuolo e che interferiscono con il normale metabolismo del microorganismo, sono esempi di una resistenza generale preformata. Spesso tale resistenza non è sufficiente a determinare l'esclusione del patogeno e diventa cruciale la velocità di riconoscimento e di attivazione della risposta [69]. È questo il caso della resistenza di tipo incompatibile, determinata dal riconoscimento del patogeno mediante interazione gene-per-gene, che porta alla risposta ipersensibile (HR) come meccanismo di difesa. L'HR comporta la morte cellulare programmata (PCD) delle cellule vegetali vicine al sito di infezione, limitando la diffusione della malattia [70]. I tempi e la sequenza degli eventi che seguono il riconoscimento gene-per-gene possono essere molto differenti a seconda del tipo di patogeno e di ospite. Essi sono tuttavia accomunati da almeno tre eventi chiave: l'aumento della concentrazione di calcio citosolico, l'aumento dell'attività chinasi e l'avvento di un "burst" ossidativo e nitrosativo [71]. Uno degli eventi più precoci che avvengono durante l'HR infatti è il rapido accumulo di ROS e di NO [67, 68]. L'incremento indipendente di una di queste due componenti ha pochi effetti sull'induzione della morte cellulare [67, 72].

La maggior parte dei dati sperimentali disponibili sulla produzione di NO durante l'interazione pianta-patogeno deriva da studi riguardanti l'infezione da parte di patogeni biotrofi [73]. Ad un immediato picco di produzione di NO, che raggiunge il suo massimo circa 1 h dall'infezione con un ceppo avirulento di *Pseudomonas*

syringae, in cellule di tabacco e soia, segue un secondo picco compreso tra 4 e 8 h dopo il trattamento [67, 68]. Un accumulo di NO è stato osservato anche in foglie di tabacco trattate con l'elicettore fungino criptogeina [74] e in tuberi di patata trattati con elicitori di *Phytophthora infestans* [75].

Un'ulteriore dimostrazione dell'importanza del ruolo dell'NO nell'HR è data dallo studio di mutanti *Atnos1* di *A. thaliana* durante l'interazione con una razza avirulenta di *P. syringae*. Questi presentano una maggior crescita del patogeno e l'insorgenza più rapida e più severa dei sintomi della malattia rispetto al "wild type" [76]. Come NO e ROS cooperino nell'induzione della morte cellulare ipersensibile non è ancora noto, ma è chiaro che questo processo dipende fortemente dal loro equilibrio (Fig. 1.8).

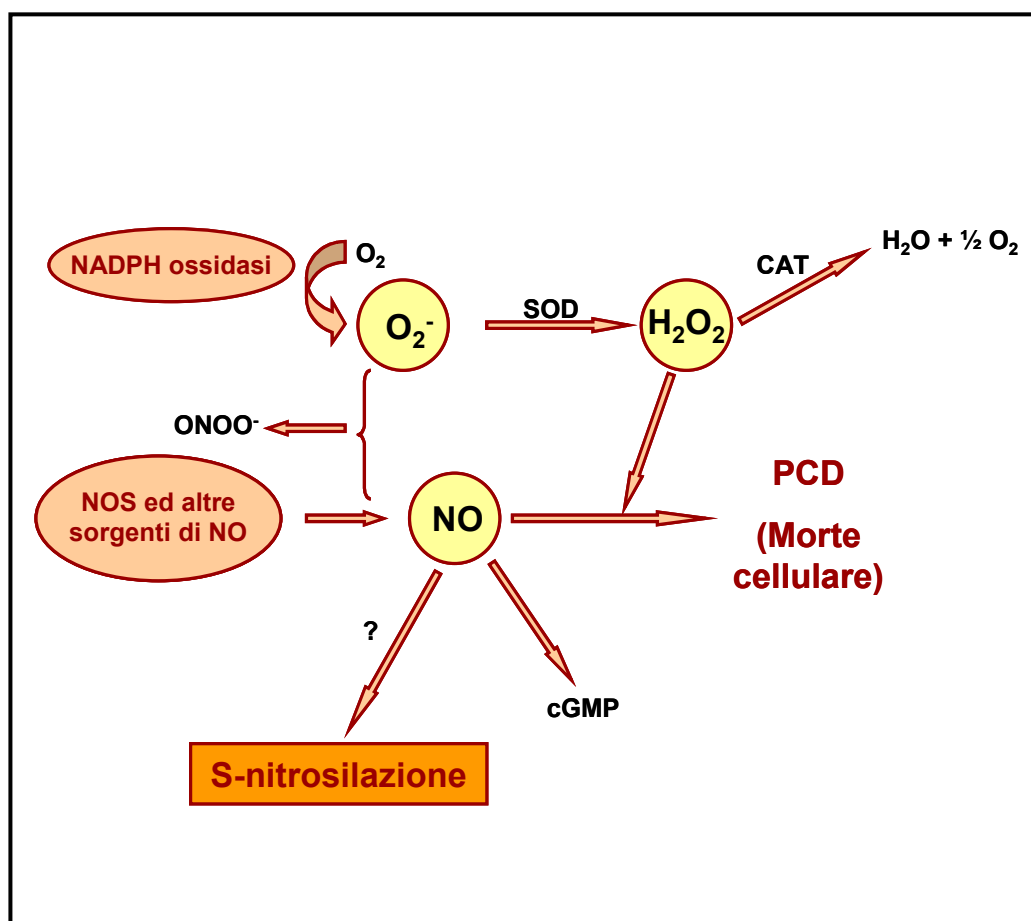


Figura 1.8 Modello di azione dell'NO e dell'H₂O₂ nella risposta ipersensibile.

Inoltre la PCD indotta da NO richiede un metabolismo attivo della cellula che comporta il processamento dell'RNA e la sintesi di nuove proteine [77]. È stato dimostrato che viene attivato un processo in cui le cistein proteasi svolgono

un ruolo chiave [78] e geni codificanti cistein proteasi, sono stati trovati essere indotti da NO in *A. thaliana* [79]. Un ruolo fondamentale è svolto dal cGMP, noto messaggero secondario dell'NO in sistemi animali che può regolare direttamente o indirettamente diverse funzioni cellulari modulando l'attività di molteplici canali del calcio [80]. L'NO regola la mobilitazione di Ca^{2+} intracellulare [81] e influisce sull'espressione di numerosi geni di difesa [82].

CAPITOLO 2

MATERIALI E METODI

2.1 Materiale vegetale

Nel presente lavoro di tesi sono state utilizzate piante di *Arabidopsis thaliana*, ecotipo Columbia 0. Le piante sono fatte crescere in condizioni di giorno breve (8 ore di luce; 16 ore di buio), mantenendo un'umidità relativa del 60%, una temperatura diurna di 25°C e notturna di 22°C.

2.2 Materiale microbiologico

Nello studio sulla variazione della S-nitrosilazione durante l'interazione pianta-patogeno è stato utilizzato il ceppo avirulento di *Pseudomonas syringae* pv *tomato* DC3000 contenente il plasmide di avirulenza *AvrB* riconosciuto dal gene di resistenza RPM1 presente nell'ecotipo Columbia 0 di *A. thaliana* (*Pst AvrB*). I batteri sono stati incubati a 28°C in KB ("King's B") contenente gli antibiotici Kanamicina (50 mg/ml) e Rifampicina (50 mg/ml). Il KB contiene peptone 2% (w/v), glicerolo 1% (v/v), K₂HPO₄ 8,6 mM e MgSO₄ 6,1 mM a pH 7,2.

2.3 Infiltrazione di foglie di *A. thaliana* con *Pst AvrB*

Una colonia di *Pst AvrB* viene posta a crescere o/n in 5 ml di terreno liquido adatto fino a 1 OD₆₀₀/ml. Le cellule vengono centrifugate a 4.000 g a 15° C per 1 min. e il precipitato lavato 2 volte con 1 ml di MgCl₂ 10 mM sterile e infine risospeso in 1 ml di MgCl₂ 10 mM sterile. L'OD₆₀₀ viene nuovamente ricontrollata. In base al valore ottenuto (1 OD₆₀₀/ml corrisponde a 1x10⁹ cell/ml) si prepara la diluizione per l'infiltrazione (10⁸ cell/ml).

L'infiltrazione in foglia delle cellule batteriche avviene mediante siringa da 1 ml, senza ago. 7 µl di sospensione cellulare vengono posti sulla punta della siringa, si appoggia il foro di uscita della siringa alla pagina inferiore della foglia, dove si trovano gli stomi, e si esercita una leggera pressione che consenta al liquido di penetrare.

2.4 Estrazione proteica

Le foglie di *A. thaliana* (0,2-0,6 g) vengono triturate in un mortaio utilizzando il tampone di estrazione (Hepes 25 mM pH7,7; EDTA 1 mM; TritonX-100 0,2% (v/v); inibitori delle proteasi, della Sigma, 1X) in un rapporto 1:3 (w/v). La

soluzione ottenuta viene centrifugata a 4° C per 30 min. a 13.000 rpm. Il surnatante viene recuperato e conservato a -80 °C.

2.5 Purificazione per immunoprecipitazione delle proteine S-nitrosilate

L'isolamento delle proteine S-nitrosilate dall'estratto proteico totale di *A. thaliana* viene effettuato applicando una variante del metodo "biotin switch". Tale variante prevede l'immunoprecipitazione delle proteine biotinilate, utilizzando NeutrAvidina immobilizzata (N-Av agarosio), seguita dall'eluizione delle proteine con β -mercaptoetanololo.

L'estratto proteico viene addizionato con 3 volumi di "blocker solution", contenente Hepes 25 mM pH 7,7; EDTA 1 mM; SDS 3,3 % (w/v); MMTS 27 mM da una stock 2 M diluito in dimetilformammide (DMF). Il campione viene incubato a 50 °C per 30 min. Per eliminare l'MMTS, si aggiungono 2 volumi di acetone (-20 °C), si incuba a -20 °C per 20 min e si centrifuga per lo stesso tempo a 2.500 g a 4 °C. Il precipitato così ottenuto è risospeso in tampone di HENS, Hepes 25 mM; EDTA 1 mM; SDS 1% (w/v). A questo vengono aggiunti ascorbato, alla concentrazione finale di 1 mM, e biotina-HPDP ("N-[6-(Biotinamido)hexyl]-3'-(2'-pyridyldithio)propionamide") alla concentrazione finale di 1,1 mM. I campioni sono poi incubati al buio per 1 h e 30 min. a temperatura ambiente per consentire la completa biotinilazione. A seguito dell'incubazione, le proteine vengono nuovamente precipitate tramite il trattamento con acetone e risospese in tampone HENS aggiungendo 2 volumi di tampone di neutralizzazione (Hepes 25 mM pH 7,7; EDTA 1 mM; NaCl 100 mM; Triton X-100 0,01 mM) e 30 μ l di soluzione di N-Av agarosio per ogni mg di proteina di partenza. Per consentire l'ottimizzazione del legame della N-Av alle proteine biotinilate i campioni vengono poi incubati per 1 h a temperatura ambiente, o/n a 4 °C e nuovamente 1h a temperatura ambiente. Dopo una centrifugazione di 5 sec. seguita dall'eliminazione del surnatante, vengono effettuati tre lavaggi con la soluzione di lavaggio (Hepes 25 mM; pH 7,7; EDTA 1 mM; 0,01 mM di TritonX 100; NaCl 600 mM), al fine di eliminare le proteine non biotinilate che non si sono legate alla N-Av. Successivamente vengono aggiunti 150 mM di tampone di eluizione (Hepes 25 mM pH 8,5; EDTA 1 mM; β -mercaptoetanololo 150 mM; NaCl 100 mM) ed incubati per 1h a temperatura ambiente, al buio, in rotazione, e per 30 min. a 37 °C. Il campione viene poi centrifugato per 5 sec. ed il surnatante, contenente le proteine che erano S-nitrosilate, viene recuperato. Segue la precipitazione con 8 volumi di acetone

(-20 °C) e 1 volume di metanolo (-20 °C), incubando per 1 h e 30 min. a -20 °C e poi centrifugando a 10.000 rpm per 30 min., a 4 °C. Il precipitato viene quindi risospeso in una soluzione di urea 7 M, tiourea 2 M, CHAPS 3% e Tris 40 mM per le analisi 2DE.

2.6 Elettroforesi bidimensionale (2-DE)

I campioni vengono solubilizzati in una soluzione contenente urea 7 M, tiourea 2 M, CHAPS 3 % e Tris 40 mM. Successivamente tale soluzione proteica viene incubata con 5 mM di trisbutilfosfina (TBP) per 1 h e 30 min. Successivamente al campione viene aggiunta iodoacetamide (IAA) alla concentrazione finale di 20 mM ed anfoliti trasportatori 0,5 %, quindi si incuba per 1 h e 30 min. Si sono utilizzate strisce IPG commerciali (Bio-Rad) da 17 cm con gradiente di pH lineare 4-7. Le strisce sono sottoposte ad una crescita lineare del voltaggio fino a 5.000V, tale voltaggio viene mantenuto o/n dopodiché viene aumentato in modo rapido a 10.000 V; questo viene mantenuto fino al raggiungimento di 75.000 Vh totali. Si è inoltre impostata una corrente limite di 50 μ A/gel e una temperatura di 20 °C.

Prima di effettuare la seconda dimensione le strisce IPG, al termine dell'isoelettrofocalizzazione, vengono incubate in una soluzione Tris-HCl 375 mM pH 8,8; urea 6 M, SDS 2% (w/v), glicerolo 20% (v/v) per 27 min. esatti, in lenta agitazione. Le strisce IPG vengono poste su gel di 20 cm di altezza, 18 cm di larghezza e 1 mm di spessore. I gel hanno un gradiente di porosità compreso tra 10% e 18% T. Il tampone di corsa utilizzato è il medesimo in entrambe le camere elettrodeiche Tris-HCl 15 mM pH 8,3; Glicina 192 mM; SDS 0,1% (w/v). Le condizioni di corsa adottate sono: 2 mA per gel 2 h, 5 mA per gel 2 h, 10 mA per gel 12 h e 20 mA per gel fino a fine corsa (18 °C). Al termine della corsa i gel vengono colorati utilizzando il colorante fluorescente Sypro Ruby o alternativamente il Coomassie colloidale.

2.7 Acquisizione ed analisi delle immagini

Le immagini dei gel colorati in SYPRO Ruby sono state digitalizzate per mezzo dello strumento "Versa-Doc" (Bio-Rad). Tutti i gel ottenuti sono stati sottoposti a scansione utilizzando il programma "Quantity-One" (Bio-Rad) che permette la selezione della risoluzione desiderata per le immagini. Tale programma, interfacciato con lo strumento "Versa-Doc", permette di effettuare

un'anteprima della digitalizzazione e solo successivamente ne acquisisce l'immagine. I tempi di acquisizione sono determinati empiricamente, generalmente nell'intorno di qualche secondo o pochi minuti. Lo strumento è in grado anche di evidenziare gli spot giunti a saturazione che devono essere il minor numero possibile, infatti questi non vengono quantificati durante l'analisi. Le immagini dei gel impiegati nel confronto quantitativo vengono filtrate e manipolate "in silico" utilizzando il software "PDQuest" (Bio-Rad). In questo modo vengono uniformate le dimensioni delle diverse mappe e viene ridotto, per quanto possibile, il rumore di fondo. La funzione "Spot Detection Wizard" rende automatico il processo di ottimizzazione dei parametri per la determinazione degli spot. Poiché le mappe variano in intensità e rumore di fondo è necessario impostare diversi parametri (sensibilità, valore minimo di picco) a seconda dei diversi tipi di gel. Le aree ad elevata densità ottica, basandosi sui parametri impostati, vengono quindi identificate come zone proteiche.

Durante la determinazione degli spot vengono create due immagini non reali: una detta immagine filtrata e una detta immagine gaussiana. La prima è l'immagine reale ripulita dai rumori di fondo, mentre la seconda è costituita dalle funzioni gaussiane associate ad ogni spot. In questo modo ciascuno spot assume una struttura tridimensionale (X, Y, Z: dove X e Y identificano la posizione sul gel della macchia e Z l'intensità del segnale in unità di densità ottica, OD) sulla quale vengono condotte tutte le analisi successive. Dopo aver selezionato il gel più rappresentativo, in termini di numero di spot e di rumore di fondo, si è proceduto alla revisione e all'allineamento degli spot, con la creazione di un gruppo di confronto ("match-set"). Il "match-set" è caratterizzato da un'immagine standard ("mappa sintetica gaussiana") che racchiude tutte le informazioni corrispondenti ad ogni macchia presente in ciascun gel e permette di individuare le corrispondenze esistenti fra le mappe. Tutte le repliche dei campioni analizzati vengono confrontate all'interno di un unico "match-set". Per compensare variazioni nell'intensità delle macchie dovute a fattori sperimentali (perdita di parte del campione durante il caricamento, tempo di colorazione insufficiente) l'intensità di ciascuno spot viene normalizzata in base all'intensità totale delle proteine presenti nel gel. Il "PDQuest" permette di condurre analisi di tipo quantitativo e statistico sugli spot presenti nei gel. L'analisi statistica effettuata prevede una variazione nella quantità proteica di almeno il 100% (è stato impostato un valore soglia = 2).

Questo valore permette di individuare differenze dovute effettivamente a variazioni nella quantità di proteine S-nitrosilate nei diversi campioni analizzati.

2.8 Digestione triptica delle proteine in-gel

Gli spot colorati con Blue Coomassie sono stati ritagliati dal gel usando un sottile bisturi e soggetti a digestione triptica in accordo con il metodo proposto da Shevchenko et al. [83]. I pezzi di gel, posti a 4° C, sono stati lasciati imbibire nel tampone di digestione contenente NH_4HCO_3 50 mM e 12,5 ng/ μl di tripsina (modified porcine trypsin, sequencing grade; Promega, Madison, WI, USA). Dopo 30 minuti l'eccesso di liquido è stato rimosso e sono stati aggiunti 20 μl di NH_4HCO_3 50 mM lasciando i frammenti di gel ad incubare overnight ad una temperatura costante di 37 °C. Al termine dell'incubazione è stato recuperato il sopranatante contenente i peptidi triptici; l'estratto, portato a secco centrifugando sotto vuoto, è stato conservato a -20 °C fino al momento dell'uso. Prima dell'utilizzo delle miscele peptidiche per le analisi di spettrometria di massa queste sono state disciolte in 10 μl FA 5% (acido formico).

2.9 Spettrometria di massa MALDI-TOF/TOF

Il metodo descritto da Gobom et al. [84] è stato utilizzato per eliminare i contaminanti salini e concentrare i peptidi triptici. Usando dei puntali GELoader (Eppendorf, Amburgo, Germania) preventivamente schiacciati all'estremità si sono impaccate delle nanocolonne dal diametro di 5mm con la resina Poros R2 (PerSeptive Biosystems, Framingham, MA, USA). Attraverso l'ausilio di una siringa che ha permesso di applicare una lieve pressione le soluzioni liquide sono state sospinte attraverso le colonne. Le colonne sono state dapprima pulite con 20 μl 100% ACN, equilibrate con 20 μl di acido formico (FA) al 5%, poi caricate con le miscele peptidiche ed infine lavate con 20 μl FA 5%. L'eluizione dei peptidi dalla colonna è stata realizzata direttamente sul target MALDI utilizzando 0.5 μl di soluzione CHCA (acido α -ciano-4-idrossicinnamico 20 $\mu\text{g}/\mu\text{l}$ in ACN/0.1% TFA v/v, 70:30 v/v). I campioni sono stati analizzati mediante lo strumento "Applied Biosystems 4700 Proteomics Analyzer MALDI-TOF/TOF" (Applied Biosystems, Framingham, MA, USA). Gli ioni positivi sono stati esaminati in modalità reflectron ed il gas utilizzato per il CID nella cella di collisione è stato aria atmosferica. L'analisi dei risultati è stata fatta utilizzando il software Data Explorer versione 4.5 (Applied Biosystem, Foster City, CA). La calibrazione degli spettri è stata realizzata internamente tramite l'utilizzo dei segnali relativi ai prodotti dell'autodigestione della

tripsina (m/z 842,51 e 2.211,10). Quando questo non è stato possibile si è ricorsi ad una calibrazione esterna con "Peptide Calibration Standard II" Bruker-Daltonics (Brema, Germania).

Per l'identificazione delle proteine è stato utilizzato il software MASCOT (<http://www.matrixscience.com>) che ha permesso di eseguire le ricerche in un database non-ridondante contenente sequenze proteiche (NCBI nr database). I seguenti parametri sono stati utilizzati per compiere le ricerche nel database: accuratezza nella determinazione della massa monoisotopica del peptide pari a 70 ppm, tolleranza sulla massa dei frammenti MS/MS $\pm 0,6$ Da, numero dei possibili siti di taglio della tripsina non utilizzati pari ad uno, completa carbamidometilazione delle cisteine e parziale ossidazione delle metionine.

CAPITOLO 3

RISULTATI

Per rilevare la presenza di proteine S-nitrosilate in *Arabidopsis thaliana*, estratti fogliari di proteine totali sono stati analizzati con una strategia basata sulla modificazione del metodo “biotin switch”. Questo approccio prevede la specifica conversione delle cisteine S-nitrosilate in cisteine biotinilate, dopo il blocco selettivo delle cisteine non nitrosilate, e la successiva immunorilevazione con anticorpi anti-biotina. Il protocollo utilizzato in questo studio è stato modificato per consentire l'identificazione e la caratterizzazione delle proteine S-nitrosilate. Tali modifiche hanno comportato l'immunoprecipitazione delle proteine biotinilate, utilizzando NeutrAvidina agarosio (N-Av) seguita dalla loro specifica eluizione con β -mercaptoetanololo. Le proteine così isolate sono state utilizzate per l'analisi del proteoma nitrosilato attraverso l'ausilio dell'elettroforesi bidimensionale (2-DE).

Questo studio è stato condotto per validare *in vivo* il protocollo di isolamento delle proteine S-nitrosilate e verificare se questa modificazione post-traduzionale è implicata, nelle piante, in processi in cui l'NO svolge un ruolo importante, quali la risposta di resistenza ai patogeni.

Foglie di *A. thaliana* sono state infiltrate con 10^8 cellule/ml di *Pseudomonas syringae* pv *tomato* che porta il gene di avirulenza *AvrB* (*Pst AvrB*) riconosciuto dal gene di resistenza RPM1 di *A. thaliana*. Campioni di foglie prelevati a 0, 4 e 8 h dall'infiltrazione sono stati sottoposti alla metodica “biotin switch” utilizzando una quantità di proteine di partenza di 3,5 mg. Sono state prodotte quattro mappe bidimensionali per ogni momento analizzato, valore ritenuto sufficiente a garantire un'accettabile significatività statistica. I gel sono stati colorati con “SYPRO Ruby” e le immagini sono state digitalizzate per mezzo dello strumento “Versa-Doc” (Bio-Rad). Tutti i gel ottenuti sono stati sottoposti a scansione utilizzando il programma “Quantity-One” (Bio-Rad; Fig. 3.1) .

Le immagini delle mappe bidimensionali sono state filtrate e manipolate “in silico” per mezzo del software “PD-Quest” (Bio-Rad), presso il laboratorio del Prof. Righetti. Tale programma permette di effettuare l'analisi quantitativa degli spot nei diversi gel. Dopo aver selezionato il gel più rappresentativo, in termini di numero di spot e di rumore di fondo, si è proceduto con la creazione di un gruppo di confronto (“match-set”). Questo contiene la mappa standard (Fig. 3.2) e

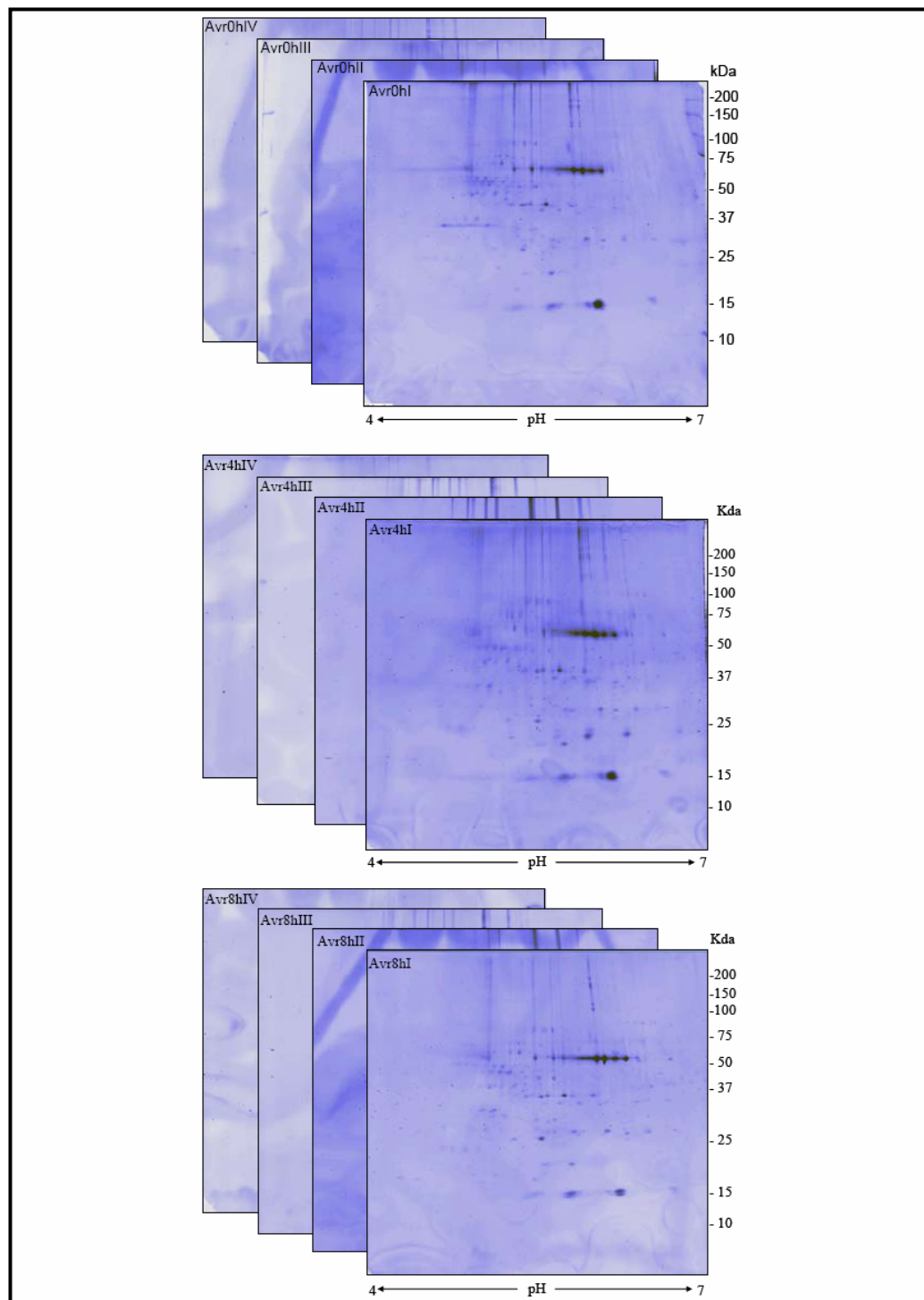


Figura 3.1 Mappe bidimensionali di proteine S-nitrosilate da campioni di foglie di *A. thaliana* prelevati a (1) 0 h; (2) 4 h; (3) 8 h dall'infiltrazione con *Pst AvrB*. Strisce 17cm, pH 4-7; SDS-PAGE con gradiente 10-18% T, colorazione tramite "SYPRO-Ruby stain".

tutte le repliche delle immagini gaussiane (vedi materiali e metodi) dei gel appartenenti ai diversi gruppi sperimentali.

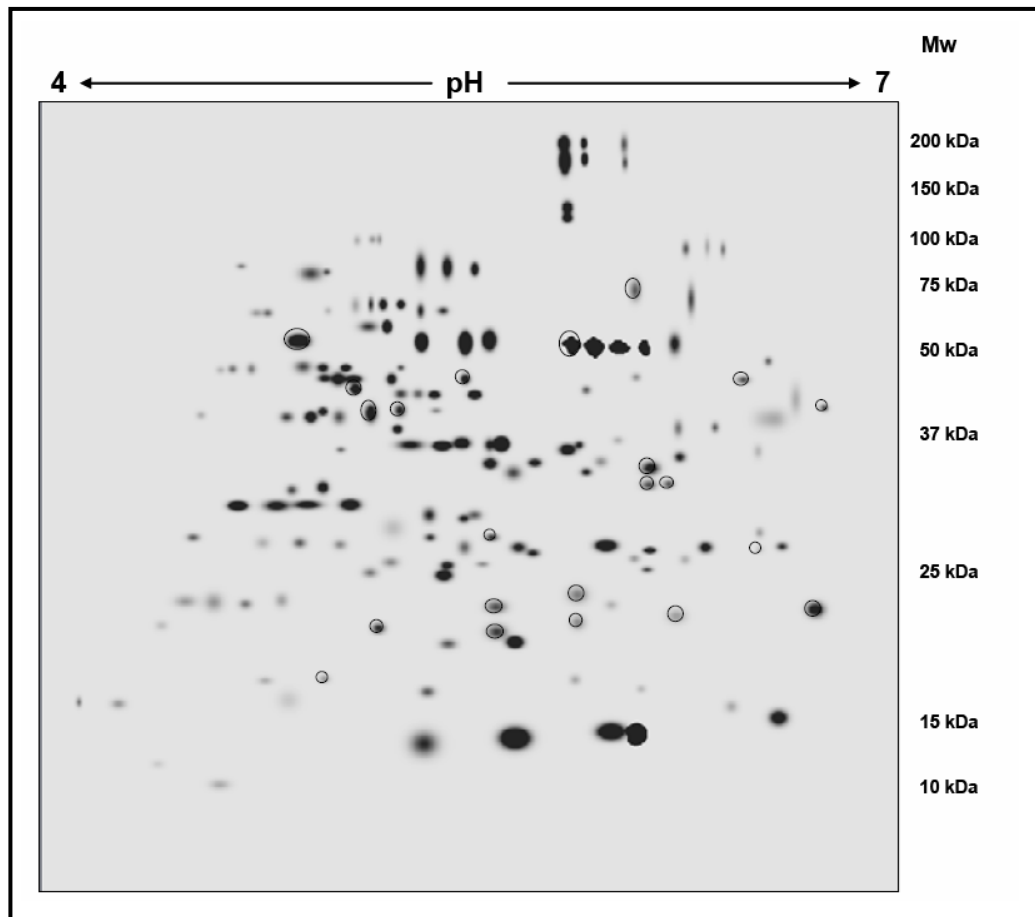


Figura 3.2 Mappa standard dei campioni analizzati.

Dopo aver allineato tutte le mappe 2D tra loro è stata normalizzata l'intensità di ogni spot in base all'intensità totale degli spot presenti in ogni gel. Quindi, le analisi quantitative e l'analisi statistica sono state effettuate allo scopo di individuare gli spot proteici diversamente espressi in modo significativo dopo l'infiltrazione con il patogeno. Per le analisi quantitative il valore soglia selezionato è stato 2,0 X in modo da individuare le proteine diversamente espresse almeno del 100%, mentre per il test statistico t di Student è stato selezionato un valore di $\alpha = 0,05$. Comparando le mappe delle proteine S-nitrosilate di campioni prelevati a 0 h (ovvero il gruppo di appartenenza 0 h) con quelle di campioni prelevati a 4 h dall'infiltrazione con *Pst AvrB*, è stato possibile identificare 12 spot che variavano in modo statisticamente significativo (Fig. 3.3).

Nella figura è stata rappresentata, attraverso degli istogrammi, la quantità di proteina nei diversi gruppi di appartenenza.

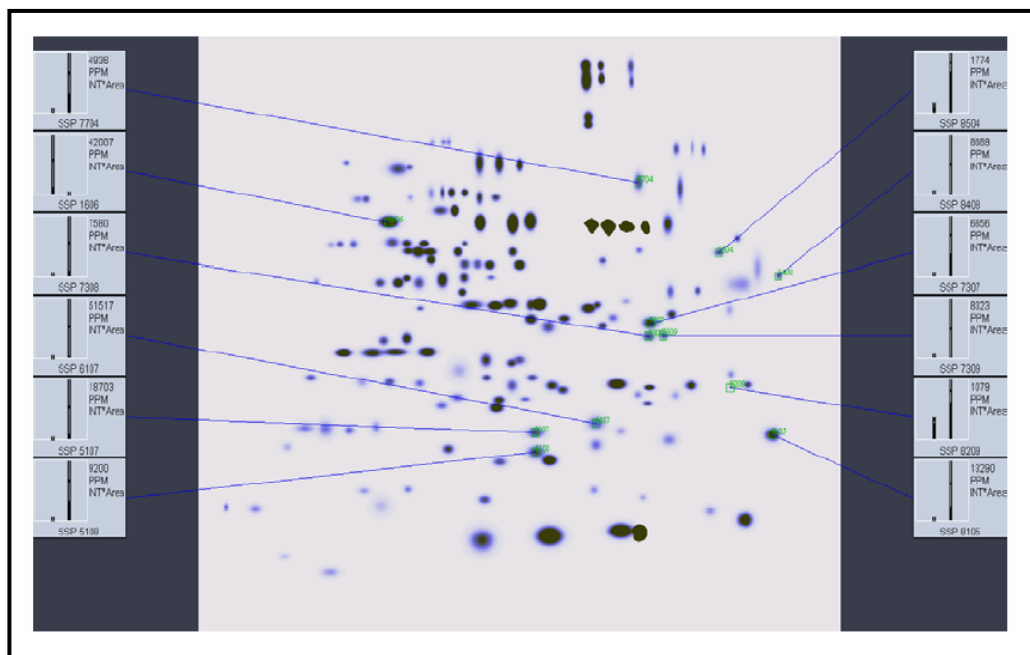


Figura 3.3 Spot variati in modo statisticamente significativo confrontando i gruppi di appartenenza 0 h e 4 h dall'infiltrazione con *Pst AvrB*. Gli istogrammi rappresentano la quantità media ed i rispettivi coefficienti di variazione di uno spot proteico nei due diversi gruppi di appartenenza.

Confrontando il gruppo di appartenenza 0 h e 8 h dall'infiltrazione con *Pst AvrB* sono stati trovati 4 spot che variavano significativamente (Fig. 3.4).

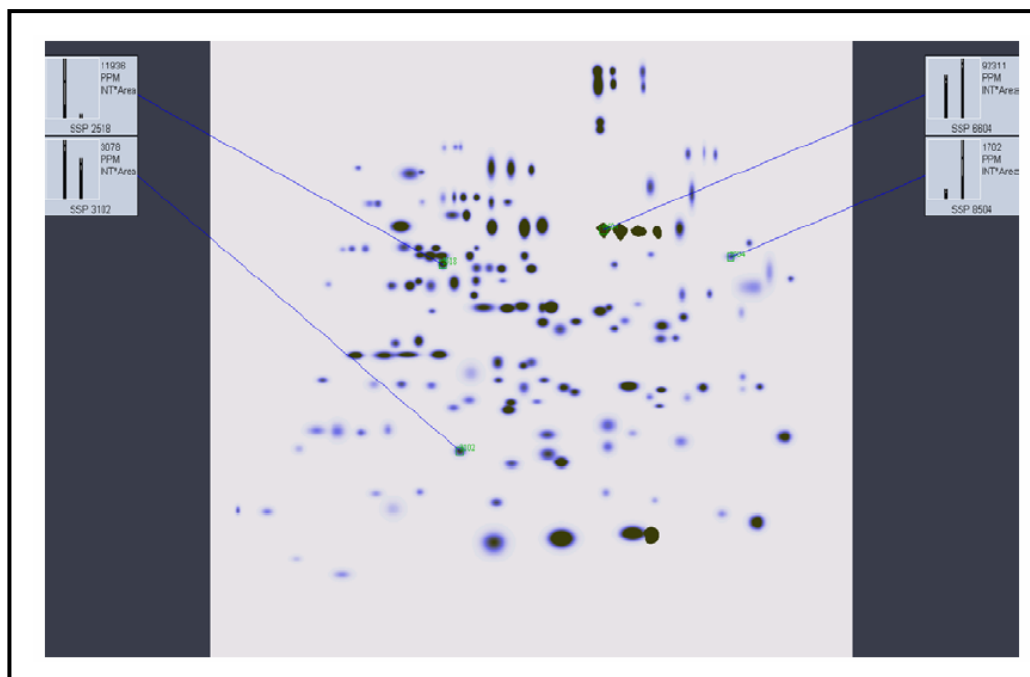


Figura 3.4 Spot variati in modo statisticamente significativo confrontando i gruppi di appartenenza 0 h e 8 h dall'infiltrazione con *Pst AvrB*. Gli istogrammi rappresentano la quantità media ed i rispettivi coefficienti di variazione di uno spot proteico nei due diversi gruppi di appartenenza.

Infine sono stati trovati 14 spot differenti confrontando i campioni a 4 h e a 8 h dopo l'infiltrazione (Fig. 3.5).

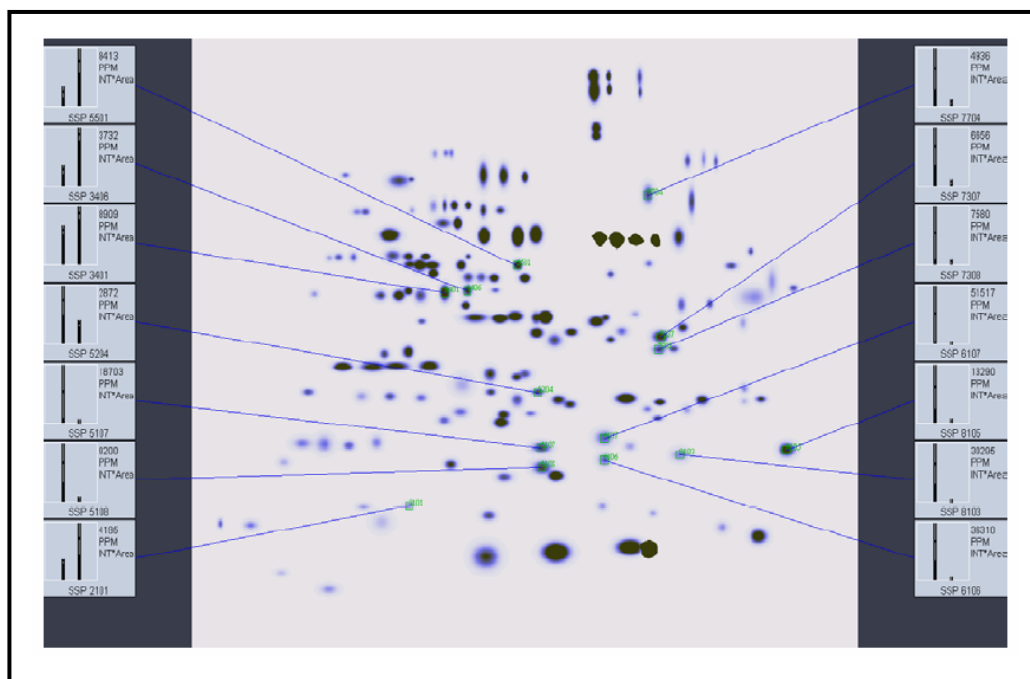


Figura 3.5 Spot variati in modo statisticamente significativo confrontando i gruppi di appartenenza 4 h e 8 h dall'infiltrazione con *Pst AvrB*. Gli istogrammi rappresentano la quantità media ed i rispettivi coefficienti di variazione di uno spot proteico nei due diversi gruppi di appartenenza.

In totale sono stati individuati 22 spot differenti modulati durante i vari trattamenti (Tabella 3.1). Alcuni di questi variavano in diversi gruppi di confronto: lo spot SSP 8504 variava sia confrontando le 0 h con le 4 h che confrontando le 0 con le 8 h; gli spot SSP 5108, 7307, 8105, 7308 e 6107 variavano sia nelle 0 h vs 4 h che nelle 4 h vs 8 h.

Gli spot sono stati escissi digeriti con tripsina ed analizzati con l'ausilio di uno spettrometro di massa MALDI-TOF/TOF. Questo ha permesso di ottenere, non solo informazioni relative al "peptid mass fingerprinting" (PMF) delle proteine incognite, ma anche informazioni di sequenza attraverso esperimenti di MS/MS. Molti campioni sono risultati non sufficientemente concentrati per consentire l'analisi MS. Per avere la sicurezza che la quantità di proteine presente negli spot prelevati fosse compresa nell'intervallo di rilevabilità dello strumento, è stato fissato come limite minimo la visibilità della macchia utilizzando la colorazione "Coomassie colloidale". Per tale ragione nuove mappe bidimensionali sono state prodotte utilizzando 21 mg (anziché 3,5 mg) di proteine di partenza.

<i>ID spot</i>	0h	4h	8h
1606	++	-	+
2101	+++	+	+++
2518	++	-	+
3102	+++	+	+++
3401	+++	+	+++
3406	+++	+	+++
5107	-	+++	-
5108	-	+++	-
5204	+++	+++	+
5501	+++	+	+++
6106	+	++	-
6107	-	+++	-
6604	+++	+	+++
7307	-	+++	-
7308	-	+++	-
7309	-	++	+
7704	-	+++	-
8103	+	++	-
8105	-	+++	-
8209	+	+++	+++
8408	-	++	+
8504	-	+++	+++

Tabella 3.1 Spot modulati durante i vari trattamenti. (*) le differenze trovate in questi trattamenti non sono statisticamente significative.

I gel sono stati colorati con “Coomassie colloidale” e gli spot non identificati precedentemente sono stati rintracciati mediante l’allineamento con la mappa standard e sono state utilizzati per l’identificazione. L’aumento della

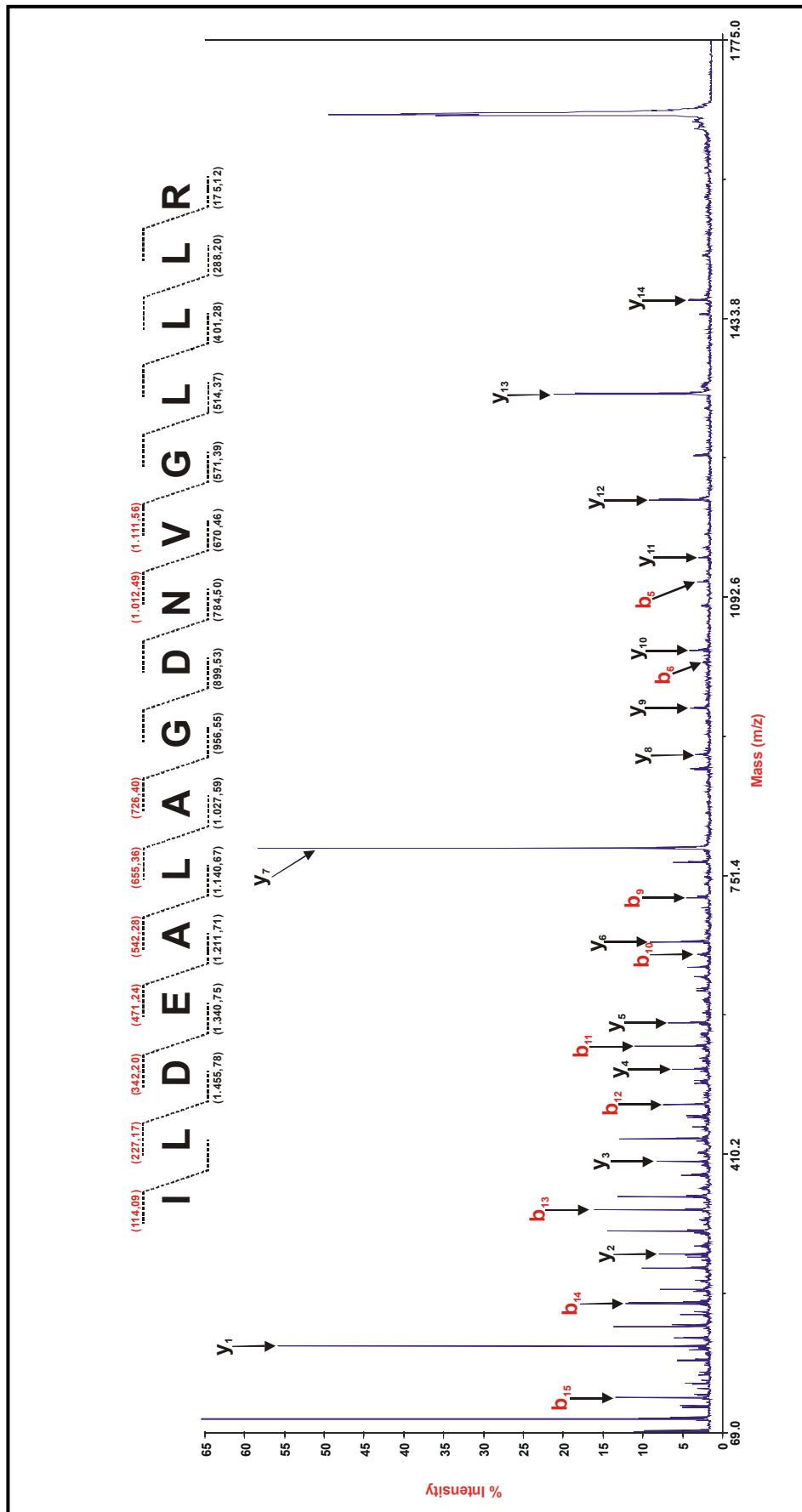


Figura 3.6 Spettro MS/MS relativo al peptide 1.691,85 di sequenza ILDEAGDNVGLLLR ottenuto dalla digestione dello spot 5501 identificato come “putative translation elongation factor EF-Tu precursor, chloroplast”.

concentrazione proteica utilizzata per la separazione bidimensionale ha permesso di avere materiale a sufficienza per ottenere ottimi spettri MS/MS (Fig. 3.6) garantendo l'identificazione inequivocabile dei prodotti proteici con elevate coperture di sequenza. I risultati di questa indagine sono riassunti in Tabella 3.2.

Spot	Numero di peptidi	NCBI Accession number	Peptidi Sequeunziati (ion score)
Proteina	Copertura di sequenza	PM pI	
1606			
large subunit of ribulose 1,5-bisphosphate carboxylase/oxygenase	25	gi 5881702	DTDILAAFR (42) VALEACVQAR (37) DNGLLLHIHR (33) DLAVEGNEIR (47) TFQGPPHGIQVER (79) EITFNFTIDKLDGQE (62)
	48%	53435	
		5.88	
2101			
putative peroxiredoxin	12	gi 23297326	RYAILADDGVVK (42) TILFAVPGAFTPTCSQK (113) VLNLEEGGAFTNSSAEDM(ox)LK (132)
	49%	24783	
		9.12	
2518			
putative phosphoglycerate kinase	18	gi 17065610	ADLNVPLDDNQITDDTR (148) FSLAPLVPR (60) LASLADLYVNDAFGTAHR (99) ELDYLVGAVSNPK (106) RPFAAIVGGSK (41) IVPASGIEDGWM(ox)GLDIGPDSIK (57)
	47%	50195	
		5.91	
3102			
large subunit of ribulose 1,5-bisphosphate carboxylase/oxygenase	10	gi 5881702	LTYTPEYETK (39) LTYTPEYETKDTDILAAFR (95) DTDILAAFR (55) ALAAALRLEDLR (48) TFQGPPHGIQVER (69)
	26%	53435	
		5.88	
3401			
phosphoribulokinase precursor	18	gi 23197622	LTSVFGGAAPPK (65) FYGEVTQQMLK (52) KPDFDAFIDPQK (100) ILVIEGLHPM(ox)FDER (73)
	51%	44721	
		5.71	
3406			
phosphoribulokinase precursor	18	gi 23197622	ANDFDLM(ox)YEQVK (69) ILVIEGLHPM(ox)FDER (39) FYGEVTQQM(ox)LK (56) IRDLYEQLIANK (52)
	60%	44721	
		5.71	
5107			
allene oxide cyclase	6	gi 34391982	ALSQNGNIENPRPSK (36) VQELSVYEINELDR (61) VQELSVYEINELDRHSPK (42)
	33%	27732	
		6.90	
5108	X	X	X
5204			
cytosolic triosephosphate isomerase	14	gi 7076787	FFVGGNWK (54) VAYALAQGLK (80) VIACVGETLEER (68) AILNESSEFVGDK (85) SDFVAAQNCWVK (48) VASPAQAQEVHDELK (60) VTNWSNVVIAIEPVWAIGTK (69) ELGGQADVDGFLVGGASLKPEFIDIK (120)
	57%	27380	
		5.39	

Tabella 3.2 Identificazioni delle proteine nitrosilate differenzialmente espresse.

Spot	Numero di peptidi	NCBI Accession number	Peptidi Sequeunziati (ion score)
Proteina	Copertura di sequenza	PM pI	
5501 putative translation elongation factor EF-Tu precursor, chloroplast	19 44%	gi 24030503 51883 5.84	VGETVDLVGLR (59) KYDEIDAAPEER (57) ILDEALAGDNVGLLLR (137) HSPFFAGYRPQFYM(ox)R (44) IYELMDAVDDYIPIQR (73) IYELM(ox)DAVDDYIPIQR (39)
6106	X	X	X
6107 nClpP1	8 30%	gi 5360579 32506 8.35	EYGLIDGVIM(ox)NPLK (58) INQDTRDRDFM(ox)SAK (40) EAKEYGLIDGVIM(ox)NPLK (60) ANLNGYLAYHTGQSLEK (52)
6604 large subunit of ribulose 1,5-bisphosphate carboxylase/oxygenase	33 56%	gi 5881702 53435 5.88	DTDILAAFR (65) VALEACVQAR (62) DLAVEGNEIIR (69) FLFCAEAIYK (80) EITFNFTIDK (43) TFQGPPHGIQVER (102) YGRPLLGTIKPK (39) WSPELAAACEVWK (78) GHYLNATAGTCEEMIK (42) EITFNFTIDKLDGQE (73) GGLDFTKDDENVNSQPFMR (88) ELGVPIVMHDYLTGGFTANTSLSHYCR (110) ALEGADLVIIIPAGVPR (49) SEVVGVMGDDNLAK (37)
7307 putative mitochondrial NAD-dependent malate dehydrogenase	5 23%	gi 12642848 36010 8.54	
7308	X	X	X
7309 At5g35100	8 30%	gi 51536584 30771 6.82	IFPGQYFLAGR (56) QFNDFAEFLGDER (44) VFLDFSLCPTYFR (42)
7704 large subunit of ribulose 1,5-bisphosphate carboxylase/oxygenase	18 38%	gi 5881702 53435 5.88	DTDILAAFR (56) VALEACVQAR (41) DLAVEGNEIIR (59) ALAAALRLEDLR (50) TFQGPPHGIQVER (64) NEGRDLAVEGNEIIR (53) LEGDRESTLGFVDLLR (37) GGLDFTKDDENVNSQPFM(ox)R (45) AAVTPAFAPAYAGINGLGVSLAR (57) GPQSPSGYSCK (35)
8103 germin-like protein	3 28%	gi 1755188 19382 5.86	
8105 Oxygen-evolving enhancer protein 3 precursor-like protein	8 40%	gi 21593587 24628 9.72	VGGPPLPSGGLPGTDNSDQAR (149) FYIQLSPTEAAAR (112) LFQTIDNLDYAAR (113)
photosystem II oxygen-evolving complex protein 3-like	10 48%	gi 21593902 23781 9.64	FYLQPLPTEAAAR (97) DIINVKPLIDR (61)

Tabella 3.2 Identificazioni delle proteine nitrosilate differenzialmente espresse.

Spot	Numero di peptidi	NCBI Accession number	Peptidi Sequeunziati (ion score)
Proteina	Copertura di sequenza	PM pI	
8209			
glutathione S-transferase	12	gi 6056410	EPFIFR (34)
	55%	23583	KEPFIFR (45)
		6.14	VPAFEDGDFK (92) VPAFEDGDFKLFESR (97)
8408			
putative RNA binding protein	18	gi 3850621	EGHQVTLFTR (56)
	43%	42303	EAEVEPILEALPK (107)
		7.71	AGGFPEPEIVHYNPK (93) QLPGESDQDFADFSSK (112)
glyceraldehyde-3-phosphate dehydrogenase (NADP+) (phosphorylating)	16	gi 1402885	TFAEEVNAAFR (54)
	42%	42766	MASVTFSPVK (71)
		7.62	VPTPNVSVVDLVVQVSK (67)
8504			
putative monodehydroascorbate reductase protein	19	gi 22022508	EAVAPYERPALSK (52)
	50%	46629	GylfPEGAAR (72)
		6.41	GADSKNLYLR (71) NILYLR (38) AAEGGAVEEYDLPFFYSR (66) FGAYWVQGGK (40) ARPSAESLDELVK (33)

Tabella 3.2 Identificazioni delle proteine nitrosilate differenzialmente espresse.

CAPITOLO 4

DISCUSSIONE

L'ossido nitrico (NO) è una molecola che esercita numerose e diverse attività in specie anche molto distanti filogeneticamente. Un importante obiettivo nello studio dei suoi meccanismi di azione è oggi rappresentato dall'identificazione di proteine che subiscono S-nitrosilazione/denitrosilazione ad opera di specie chimiche da esso derivate. Le evidenze sperimentali, in continuo aumento in ambito animale, indicano che la S-nitrosilazione/denitrosilazione di specifici residui cisteinici delle catene polipeptidiche è una componente importante di regolazione dell'attività proteica che si configura come un analogo della fosforilazione/defosforilazione.

Il recente sviluppo di metodi che permettono di identificare residui cisteinici S-nitrosilati nelle proteine ha inoltre aperto la via all'identificazione del proteoma S-nitrosilato. Dato che la produzione e l'azione fisiologica dell'NO nelle piante è ampiamente dimostrata [1] è assolutamente probabile la S-nitrosilazione di residui cisteinici nelle proteine di *Arabidopsis thaliana* in diverse condizioni pato-fisiologiche per le quali è già noto un coinvolgimento dell'NO a livello fisiologico. A supporto di tale ipotesi recentemente è stato pubblicato uno studio bioinformatico a seguito di una ricerca in banca dati dei possibili substrati della S-nitrosilazione, sulla base della presenza di un motivo conservato caratteristico di molte proteine animali S-nitrosilate [85, 26]. Tale studio ha portato all'identificazione di 99 proteine di pianta putativamente S-nitrosilabili appartenenti a diverse classi funzionali [85]. La presenza del motivo conservato non è tuttavia necessaria né sufficiente per poter affermare che la proteina viene S-nitrosilata, dato che, per esempio, in molti sistemi biologici sono state identificate numerose proteine S-nitrosilate che ne sono prive [15].

Questo lavoro ha avuto lo scopo di identificare un metodo per la rilevazione delle proteine S-nitrosilate in vivo e dimostrare l'esistenza e la funzione di questa modificazione post-traduzionale anche nei sistemi vegetali.

Si è deciso quindi di adottare un metodo indiretto di rilevazione delle proteine S-nitrosilate, il "biotin switch"[45]. La metodica si basa, come indica il nome, sulla derivatizzazione specifica dei residui cisteinici S-nitrosilati con biotina e sul successivo riconoscimento mediato da un anticorpo anti-biotina. Per permettere l'ottenimento di un'elevata quantità di proteine tale da consentire una

successiva analisi proteomica, è stato necessario mettere a punto un protocollo ad hoc modificando la strategia originale. Per la purificazione delle proteine S-nitrosilate è stata utilizzata NeutrAvidina che permette l'immunoprecipitazione delle proteine biotinilate. Per ottenere la separazione delle proteine S-nitrosilate così purificate è stata applicata una metodica di 2D elettroforesi. L'utilizzo di questa tecnica di separazione delle proteine presenta un valido supporto all'analisi per il suo elevato grado di risoluzione. In sistemi animali tale metodo è stato utilizzato per identificare le proteine S-nitrosilate in cellule mesangiali e ha permesso la risoluzione di 790 spot; di questi solo i 34 più abbondanti sono stati analizzati ed identificati [47].

Nessuno studio è mai stato effettuato sulla S-nitrosilazione in sistemi vegetali in vivo e solo recentemente Lindermayr e collaboratori, trattando estratti cellulari di *A. thaliana* con GSNO, hanno isolato in pianta proteine S-nitrosilate in vitro [86].

In questo studio, il sistema di isolamento delle proteine S-nitrosilate è stato invece validato in vivo studiando la variazione del profilo di proteine S-nitrosilate in un processo in cui l'NO svolge un ruolo importante, quale l'interazione incompatibile pianta-patogeno. L'infezione di *A. thaliana* da parte del patogeno avirulento Pst AvrB comporta, infatti, un "burst" nitrosativo che raggiunge il suo picco massimo in un intervallo di tempo compreso tra le 4 e le 8 h dall'infiltrazione [1]. Sono state quindi purificate le proteine S-nitrosilate ottenute da campioni di foglie di *A. thaliana* prima, durante e dopo il "burst" nitrosativo. Successivamente tali proteine sono state separate attraverso 2D elettroforesi e i profili di mappa ottenuti dai diversi campioni sono stati analizzati mediante il programma PD-Quest. L'analisi statistica dei dati ha permesso di rilevare le differenze nei profili delle proteine S-nitrosilate durante l'interazione pianta-patogeno e ha portato all'individuazione di 22 spot differenti che variavano in modo significativo.

Tali spot sono stati escissi digeriti proteoliticamente con tripsina ed analizzati con uno spettrometro di massa MALDI-TOF/TOF. L'analisi in MS ha prodotto non solo informazioni relative al PMF delle proteine incognite, ma anche informazioni di sequenza grazie alla possibilità di realizzare esperimenti di tandem MS. I dati acquisiti utilizzati per l'interrogazione delle banche dati, tramite il motore di ricerca Mascot, hanno portato all'identificazione di 19 proteine; tre spot non sono stati identificati in quanto non presentavano una quantità sufficiente di proteina per consentire l'analisi (Tabella 3.2). A supporto della validità dei risultati

ottenuti con la metodica messa a punto è stato osservato che alcune delle proteine identificate nel corso di questo lavoro sono state identificate come S-nitrosilate anche in studi effettuati su estratti proteici di colture di *A. thaliana* trattati *in vitro* con GSNO [86]:

- gliceraldeide 3-P deidrogenasi (GAPDH),
- perossiredossina,
- fattori di allungamento della traduzione,
- ribulosio1,5-bisfosfato carbossilasi/ossigenasi
- oxygen-evolving enhancer
- glutatione S-transferasi (GST)
- triosofosfato isomerasi.

Inoltre proteine omologhe a quelle identificate sono state rilevate come S-nitrosilate anche in sistemi animali. Infatti, numerosi studi sia *in vitro* sia *in vivo* hanno dimostrato che la GAPDH in sistemi animali può essere S-nitrosilata sulla cisteina del sito attivo (Cys 149) [87, 88], con conseguente inibizione enzimatica. La S-nitrosilazione della GAPDH comporta inoltre il legame della proteina a Siah1, una ligasi ubiquitina E3; ciò influenza la citotossicità e l'apoptosi mediando la degradazione delle proteine nucleari [89]. Una perossiredossina omologa a quella identificata in questo studio è stata trovata S-nitrosilata in cellule endoteliali animali [48]. Tale proteina presenta due residui di cisteina molto conservati nel sito attivo soggetti a regolazione ossido riduttiva, la loro S-nitrosilazione potrebbe rappresentare un meccanismo di modulazione della funzione proteica. Studi su topo hanno evidenziato l'attivazione e l'inibizione rispettivamente di GST microsomiali e citosoliche dopo trattamento con GSNO. Ciò suggerisce che tale classe di proteine è soggetta a modulazione dell'attività da parte della S-nitrosilazione.

Anche lo studio dell'S-nitroso proteoma di *Mycobacterium tuberculosis* ha portato all'identificazione di proteine S-nitrosilate omologhe alle proteine qui identificate come le proteine appartenenti alla classe dei fattori di allungamento della traduzione [90].

Questo progetto di ricerca ha permesso di mettere a punto un sistema che consente l'identificazione in pianta di proteine S-nitrosilate *in vivo* ed ha dimostrato la presenza di tale modificazione post-traduzionale anche nei sistemi vegetali. La validità dello studio effettuato è supportata dal fatto che circa il 40% delle proteine rilevate differenzialmente S-nitrosilate sono state rinvenute S-nitrosilate anche in

altri sistemi cellulari. Questo dimostra da un lato l'efficacia della metodica applicata e dall'altro le notevoli potenzialità per l'identificazione di un gran numero di proteine S-nitrosilate.

CAPITOLO 5

BIBLIOGRAFIA

- [1] Delledonne, M. (2005) NO news is good news for plants. *Curr. Opin. Plant Biol.* **4**: 390-396.
- [2] Lamattina, L., Garcia-Mata, C., Graziano, M., Pagnussat, G. (2003) Nitric oxide: the versatility of an extensive signal molecule. *Annu. Rev. Plant Biol.* **54**: 109-136.
- [3] Mayer, B. (2000) Nitric oxide. *Handbook of Experimental Pharmacology* Springer-Verlag **143**: 667.
- [4] Fang, F. C. (1999) Nitric oxide and infection. New York: Kluwer Academic/Plenum Publishers, 475-488.
- [5] Crawford, N. M., Guo, F. Q. (2005) New insights into nitric oxide metabolism and regulatory functions. *Trends Plant Sci.* **4**: 195-200.
- [6] del Rio, L. A., Corpas, F. J., Sandalio, L. M., Palma, J. M., Barroso, J. B. (2003) Plant peroxisomes, reactive oxygen metabolism and nitric oxide. *IUBMB Life* **55**: 71-81.
- [7] Wink, D. A., Cook, J. A., Pacelli, R., Liebmann, J., Krishne, M. C., Mitchell, J. B. (1995) Nitric oxide (NO) protects against cellular damage by reactive oxygen species. *Toxicol. Lett.* **82/83**: 221-226.
- [8] Stamler, J. S., Singel, D. J., Loscalzo, J. (1992) Biochemistry of nitric oxide and its redox-activated forms. *Science* **258**: 1898-1902.
- [9] Schmidt, H. W., Walter, U. (1994) NO at work. *Cell* **78**: 919-925.

- [10] Stamler, J. S., Simon, D. I., Jaraki, O., Osborne, J. A., Francis, S., Mullins, M., Singel, D., Loscalzo, J. (1992) S-nitrosylation of tissue-type plasminogen activator confers vasodilatory and antiplatelet properties on the enzyme. *Proc. Natl. Acad. Sci. U.S.A.* **17**: 8087-8091.
- [11] Jaffrey, S. R., Snyder, S. H. (1995) Nitric oxide: a neuronal messenger. *Annu. Rev. Cell Dev. Biol.* **11**: 417-40.
- [12] Zhang, H., Shen, W. B., Zhang, W., Xu, L. L. (2005) A rapid response of B-amylase to nitric oxide but not gibberellin in wheat seeds during the early stage of germination. *Planta* **220**: 708-716.
- [13] Prado, A. M., Porterfield, D. M., Feijo, J. A. (2004). Nitric oxide is involved in growth regulation and re-orientation of pollen tubes. *Development* **11**: 2707-2714.
- [14] Beligni, M. V., Lamattina, L. (2001) Nitric oxide in plants: the history is just beginning. *Plant Cell and Environment* **24**: 267-278.
- [15] Hess, D. T., Matsumoto, A., Kim, S. O., Marshall, H. E., Stamler, J. S. (2005) Protein S-nitrosylation: purview and parameters. *Nat. Rev. Mol. Cell Biol.* **2**: 150-166.
- [16] Lucas, K. A., Pitari, G. M., Kazerounian, S., Ruiz-Stewart, I., Park, J., Schulz, S., Chepenik, P. K., Waldman, S. A. (2000) Guanylyl cyclases and signaling by cyclic GMP. *Pharmacol. Rev.* **52**: 375-414.
- [17] Martinez-Ruiz, A., Lamas, S. (2004) S-nitrosylation: a potential new paradigm in signal transduction. *Cardiovasc. Res.* **1**: 43-52.
- [18] Myers, P. R., Minor, R. L. Jr., Guerra, R. Jr., Bates, J. N., Harrison, D. G. (1990) Vasorelaxant properties of the endothelium-derived relaxing factor more closely resemble S-nitrosocysteine than nitric oxide. *Nature* **345**: 161-163.
- [19] Arnette, D. R., Stamler, J. S. (1995) NO⁺, NO and NO⁻ donation by S-nitrosylation and acceleration of disulfide formation. *Arch. Biochem. Biophys.* **318**: 279-285.

- [20] Liu, X., Miller, M. J. S., Joshi, M. S., Thomas, D. D., Lancaster, J. R. Jr. (1998) Accelerated reaction of nitric oxide with O₂ within the hydrophobic interior of biological membranes. *Proc. Natl. Acad. Sci. U.S.A.* **95**: 2175-2179.
- [21] Liu, L., Hausladen, A., Zeng, M., Que, L., Heitman, J., Stamler, J. S. (2001) A metabolic enzyme for S-nitrosothiol conserved from bacteria to humans. *Nature* **6827**: 490-494.
- [22] Liu, L., Yan, Y., Zeng, M., Zhang, J., Hanes, M. A., Ahearn, G., McMahon, T. J., Dickfeld, T., Marshall, H.E., Que, L. G., Stamler, J. S. (2004) Essential roles of S-nitrosothiols in vascular homeostasis and endotoxic shock. *Cell* **116**: 617-628.
- [23] Gow, A. J., Luchsinger, B. P., Pawloski, J. R., Singel, D. J., Stamler, J. S. (1999) The oxyhemoglobin reaction of nitric oxide. *Proc. Natl. Acad. Sci. U.S.A.* **16**: 9027-9032.
- [24] Romeo, A. A., Capobianco, J. A., English, A. M. (2003) Superoxide dismutase targets NO from GSNO to Cys β 93 of oxyhemoglobin in concentrated but not dilute solutions of the protein. *J. Am. Chem. Soc.* **125**: 14370-14378.
- [25] Lipton, S. A., Choi, Y. B., Takahashi, H., Zhang, D., Li, W., Godzik, A., Bankston, L. A. (2002). Cysteine regulation of protein function as exemplified by NMDA-receptor modulation. *Trends Neurosci.* **25**: 474-480.
- [26] Stamler, J. S., Jia, L., Eu, J. P., McMahon, T. J., Demchenko, I. T., Bonaventura, J., Gernert, K., Piantadosi, C. A. (1997) Blood flow regulation by S-nitrosohemoglobin in the physiological oxygen gradient. *Science* **276**: 2034-2037.
- [27] Hess, D. T., Matsumoto, A., Nudelman, R., Stamler, J. S. (2001) S-nitrosylation: spectrum and specificity. *Nat. Cell Biol.* **2**: E46-49.
- [28] Eu, J. P., Sun, J., Xu, L., Stamler, J. S., Meissner, G. (2000) The skeletal muscle calcium release channel: coupled O₂ sensor and NO signaling functions. *Cell* **102**: 499-509.

- [29] Lai, T. S., Hausladen, A., Slaughter, T. F., Eu, J. P., Stamler, J. S., Greenberg, C. S. (2001) Calcium regulates S-nitrosylation, denitrosylation, and activity of tissue transglutaminase. *Biochemistry* **16**: 4904-4910.
- [30] Britto, P. J., Knipling, L., Wolff, J. (2002) The local electrostatic environment determines cysteine reactivity of tubulin. *J Biol. Chem.* **277**: 29018-29027.
- [31] Perez-Mato, I., Castro, C., Ruiz, F. A., Corrales, F. J., Mato, J. M. (1999) Methionine adenosyltransferase S-nitrosylation is regulated by the basic and acidic amino acids surrounding the target thiol. *J Biol. Chem.* **274**: 17075-17079.
- [32] Nedospasov, A., Rafikov, R., Beda, N., Nudler, E. (2000) An autocatalytic mechanism of protein nitrosylation. *Proc. Natl. Acad. Sci. U.S.A.* **97**: 13543-13548.
- [33] Mannick, J. B., Schonhoff, C. M. (2002) Nitrosylation: the next phosphorylation? *Arch. Biochem. Biophys.* **1**: 1-6.
- [34] Mirza, U. A., Chait, B. T., Lander, H. M. (1995). Monitoring reactions of nitric oxide with peptides and proteins by electrospray ionization-mass spectrometry. *J. Biol. Chem.* **270**: 17185-17188.
- [35] Zech, B., Wilm, M., van Eldik, R. B. (1999) Mass spectrometric analysis of nitric oxide-modified Caspase-3. *J. Biol. Chem.* **274**: 20931-20936.
- [36] Welch, G. N., Upchurch, G. R. Jr., Loscalzo, J. (1996) S-nitrosothiol detection. *Methods Enzymol.* **268**: 293-298.
- [37] Wang, S., Wang, W., Wesley, R. A., Danner, R. L. (1999) A spl binding site of the tumor necrosis factor alpha promoter functions as a nitric oxide response element. *J. Biol. Chem.* **274**: 33190-33193.
- [38] Fang, K., Ragsdale, N. V., Carey, R. M., MacDonald, T., Gaston, B. (1998) Reductive assays for S-nitrosothiols: implications for measurements in biological systems. *Biochem. Biophys. Res. Commun.* **252**: 535-540.

- [39] Yang, B. K., Vivas, E. X., Reiter, C. D., Gladwin, M. T. (2003) Methodologies for the sensitive and specific measurement of S-nitrosothiols, iron-nitrosyls, and nitrite in biological samples. *Free. Radic. Res.* **37**: 1-10.
- [40] Wink, D. A., Kim, S., Coffin, D., Cook, J. C., Vodovotz, Y., Chistodoulou, D., Jourdain, D., Grisham, M. B. (1999) Detection of S-nitrosothiols by fluorometric and colorimetric methods. *Methods Enzymol.* **301**: 201-211.
- [41] Kostka, P., Park, J. K. (1999) Fluorometric detection of S-nitrosothiols. *Methods Enzymol.* **301**: 227-235.
- [42] Mannick, J. B., Schonhoff, C., Papeta, N., Ghafourifar, P., Szibor, M., Fang, K., Gaston, B. (2001) S-Nitrosylation of mitochondrial caspases. *J. Cell. Biol.* **6**: 1111-1116.
- [43] Schonhoff, C. M., Gaston, B., Mannick, J. B. (2003). Nitrosylation of cytochrome c during apoptosis. *J. Biol. Chem.* **278**: 18265-18270.
- [44] Matsushita, K., Morrell, C. N., Cambien, B., Yang, S. X., Yamakuchi, M., Bao, C., Hara, M. R., Quick, R. A., Cao, W., O'Rourke, B., Lowenstein, J. M., Pevsner, J., Wagner, D. D., Lowenstein, C. J. (2003) Nitric oxide regulates exocytosis by S-nitrosylation of N-ethylmaleimide-sensitive factor. *Cell* **2**: 139-150.
- [45] Jaffrey, S. R., Erdjument-Bromage, H., Ferris, C. D., Tempst, P., Snyder, S. H. (2001) Protein S-nitrosylation: a physiological signal for neuronal nitric oxide. *Nat. Cell. Biol.* **2**: 193-197.
- [46] Lopez, M. F. (1999) Proteome analysis: I. Gene products are where the biological action is. *J. Chromatogr. B* **722**: 191-202.
- [47] Kunczewicz, T., Sheta, E. A., Goldknopf, I. L., Kone, B. C. (2003) Proteomic analysis of S-nitrosylated proteins in mesangial cells. *Mol. Cell Proteomics* **2**: 156-163.

- [48] Martínez-Ruiz, A., Lamas, S. (2004) Detection and proteomic identification of S-nitrosylated proteins in endothelial cells. *Arch. Biochem. Biophys.* **423**: 192-199.
- [49] Mayer, B., Hemmens, B. (1997) Biosynthesis and action of nitric oxide in mammalian cells. *Trends Biochem. Sci.* **22**: 477-481.
- [50] Barroso, J. B., Corpas, F. J., Carrerasi, A., Sandalio, L. M., Valderrama, R., Palma, J. M., Lupiáñez, J. A., del Río, L. A. (1999). Localization of Nitric-oxide synthase in plant peroxisomes. *J. Biol. Chem.* **274**: 36729-36733.
- [51] Wendehenne, D., Pugin, A., Klessig, D. F., Durner, J. (2001) Nitric oxide: comparative synthesis and signaling in animal and plant cells. *Trends Plant Sci.* **4**: 177-183.
- [52] Modolo, L. V., Cuhna, F. Q., Braga, M. R., Salgado, I. (2002) Nitric oxide synthase-mediated phytoalexin accumulation in soybean cotyledons in response to the *Diaporthe phaseolorum f.sp. meridionalis* elicitor. *Plant Physiol.* **130**: 1288-1297.
- [53] Neill, S. J., Desikan, R., Hancock, J. T. (2003). Nitric oxide signaling in plant. *New Phytol.* **159**: 11-35.
- [54] Butt, Y. K.-C., Lum, J. H.-K., Lo, S. C.-L. (2003) Proteomic identification of plant proteins probed by mammalian nitric oxide synthase antibodies. *Planta* **216**: 762-771.
- [55] Zeidler, D., Zahringer, U., Gerber, I., Dubery, I., Hartung, T., Bors, W., Hutzler, P., Durner, J. (2004). Innate immunity in *Arabidopsis thaliana*: Lipopolysaccharides activate nitric oxide synthase (NOS) and induce defense genes. *Proc. Natl. Acad. Sci. U.S.A.* **44**: 15811-15816.
- [56] Dean, J. V., Harper, J. E. (1988) The conversion of nitrite to nitrogen oxide(s) by the constitutive NAD(P)H-nitrate reductase enzyme from soybean. *Plant Physiol.* **88**: 389-395.

- [57] Yamasaki, H., Sakihama, Y., Takahashi, S. (1999). An alternative pathway for nitric oxide production in plants: new features of an old enzyme. *Trends Plant Sci.* **4**: 128-129.
- [58] Huber, T. A., Rossmagel, P., Woese, C. R., Rachel, R., Langworthy Stetter, K. O. (1996) Formation of ammonium from nitrate during chemolithoautotrophic growth of the extremely thermophilic bacterium *Ammonifex degensii* gen. nov. sp. Nov. Syst. Appl. Microbiol. **19**: 40-49.
- [59] Yamasaki, H., Sakihama, Y. (2000) Simultaneous production of nitric oxide and peroxynitrite by plant nitrate reductase: in vitro evidence for the NR-dependent formation of active nitrogen species. *FEBS Lett.* **1**: 89-92.
- [60] Desikan, R., Griffiths, R., Hancock, J., Neill, S. (2002) A new role for an old enzyme: nitrate reductase-mediated nitric oxide generation is required for abscisic acid-induced stomatal closure in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* **99**: 16314-16318.
- [61] Corpas, F. J., Barroso, J. B., Carreras, A., Quiros, M., Leon, A. M., Romero-Puertas, M. C., Esteban, F. J., Valderrama, R., Palma, J. M., Sandalio, L. M., Gomez, M., del Rio, L. A. (2004). Cellular and subcellular localization of endogenous nitric oxide in young and senescent pea plants. *Plant Physiol.* **1**: 2722-2733.
- [62] Berthke, P. C., Badger, M. R., Jones, R. L. (2004) Apoplastic synthesis of nitric oxide by plant tissue. *Plant Cell* **16**: 332-341.
- [63] Pagnussat, G. C., Lanteri, M. L., Lombardo, M. C., Lamattina, L. (2004) Nitric oxide mediates the indole acetic acid induction activation of a mitogen-activated protein kinase cascade involved in adventitious root development. *Plant Physiol.* **1**: 279-286.
- [64] Correa-Aragunde, N., Graziano, M., Lamattina, L. (2004). Nitric oxide plays a central role in determining lateral root development in tomato. *Planta* **6**: 900-905.

- [65] Hu, X., Neill, S. J., Tang, Z., Cai, W. (2005) Nitric oxide mediates gravitropic bending in soybean roots. *Plant Physiol.* **137**: 663-670.
- [66] Zhao, L., Zhang, F., Guo, J., Yang, Y., Li, B., Zhang, L. (2004) Nitric oxide functions as a signal in salt resistance in the calluses from two ecotypes of reed. *Plant Physiol.* **134**: 849-857.
- [67] Delledonne, M., Xia, Y., Dixon, R. A., Lamb, C. (1998) Nitric oxide functions as a signal in plant disease resistance. *Nature* **6693**: 585-588.
- [68] Durner, J., Wendehenne, D., Klessig, D. F. (1998) Defense gene induction in tobacco by nitric oxide, cyclic GMP, and cyclic ADP-ribose. *Proc. Natl. Acad. Sci. U.S.A.* **17**: 10328-10333.
- [69] Bent, A. F. (1996) Plant disease resistance genes: functions meets structure. *Plant Cell* **8**: 1757-1777.
- [70] Dangl, J. (1998) Innate immunity. Plants just say NO to pathogens. *Nature* **6693**: 525, 527.
- [71] Bolwell, G. P. (1999) Role of active oxygen species and NO in plant defence. *Curr. Opin. Plant Biol.* **2**: 287-294.
- [72] de Pinto, M. C., Tomassi, F., de Gara, L. (2002) Changes in the antioxidant systems as part of the signaling pathway responsible for the programmed cell death activated by nitric oxide and reactive oxygen species in tobacco bright-yellow 2 cells. *Plant Physiol.* **130**: 698-708.
- [73] Romero-Puertas, M. C., Perazzolli, M., Zago, E. D., Delledonne, M. (2004) Nitric oxide signalling functions in plant-pathogen interactions. *Cell Microbiol.* **9**: 795-803.
- [74] Foissner, I., Wendehenne, D., Langebartels, C., Durner, J. (2000) *In vivo* imaging of an elicitor-induced nitric oxide burst in tobacco. *Plant J.* **23**: 817-824.

- [75] Yamamoto, A., Katou, S., Yoshioka, H., Doke, N., Kawakita, K. (2003) Nitrate reductase, a nitric oxide-producing enzyme: induction by pathogen signals. *J. Gen. Plant Pathol.* **69**: 218–229.
- [76] Zeidler, D., Zähringer, U., Gerber, I., Dubery, I., Hartung, T., Bors, W., Hutzler, P., Durner, J. (2004) Innate immunity in *Arabidopsis thaliana*: Lipopolysaccharides activate nitric oxide synthase (NOS) and induce defense genes *Proc. Natl. Acad. Sci. U.S.A.* **101**: 15811-15816.
- [77] Clark, D., Durner, J., Navarre, D. A., Klessig, D. F. (2000) Nitric oxide inhibition of tobacco catalase and ascorbate peroxidase. *Mol. Plant Microb. Interact.* **12**: 1380-1384.
- [78] Belenghi, B., Acconcia, F., Trovato, M., Perazzolli, M., Bocedi, A., Polticelli, F., Ascenzi, P., Delledonne, M. (2003) AtCYS1, a cystatin from *Arabidopsis thaliana*, suppresses hypersensitive cell death. *Eur. J. Biochem.* **270**: 1–12.
- [79] Polverari, A., Molesini, B., Pezzotti, M., Buonauro, R., Marte, M., Delledonne, M. (2003) Nitric oxide mediated transcriptional changes in *Arabidopsis thaliana*. *Mol. Plant Microb. Interact.* **16**: 1094-1105.
- [80] Wendehenne, D., Pugin, A., Klessig, D. F., Durner, J. (2001) Nitric oxide: comparative synthesis and signaling in animal and plant cells. *Trends Plant. Sci.* **6**: 177–183.
- [81] Lamotte, O., Gould, K., Lecourieux, D., Sequeira-Legrand, A., Lebrun-Garcia, A., Durner, J., Pugin, A., Wendehenne, D. (2004) Analysis of Nitric Oxide Signaling Functions in Tobacco Cells Challenged by the Elicitor Cryptogein. *Plant Physiol.* **1**: 516-529.
- [82] Parani, M., Rudrabhatla, S., Myers, R., Weirich, H., Smith, B., Leaman, D. W., Goldman, S. L. (2004) Microarray analysis of nitric oxide responsive transcripts in *Arabidopsis*. *Plant Biotechnol. J.* **4**: 359-366.

- [83] Shevchenko A., Wilm M., Vorm O., Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* **68**: 850–858.
- [84] Gobom J., Nordhoff E., Mirgorodskaya E., Ekman R., Roepstorff P. (1999) Sample purification and preparation technique based on nano-scale reversed-phase columns for the sensitive analysis of complex peptide mixtures by matrix-assisted laser desorption/ionization mass spectrometry. *J. Mass Spectrom.* **34**: 105–116.
- [85] Huber, S. C., Hardin, S. C. (2004) Numerous posttranslational modifications provide opportunities for the intricate regulation of metabolic enzymes at multiple levels. *Curr. Opin. Plant Biol.* **3**: 318-322.
- [86] Lindermayr, C., Saalbach, G., Durner, J. (2005) Proteomic Identification of S-Nitrosylated Proteins in *Arabidopsis*. *Plant Physiol.* **137**: 921-930.
- [87] Padgett, C. M., Whorton, A. R. (1995) S-nitrosoglutathione reversibly inhibits GAPDH by S-nitrosylation. *J. Physiol. Cell Physiol.* **269**: 739-749.
- [88] Mohr, S., Stamler, J. S., Brune, B. (1996) Posttranslational modification of glyceraldehyde-3-phosphate dehydrogenase by S-nitrosylation and subsequent NADH attachment. *J. Biol. Chem.* **271**: 4209-4214.
- [89] Hara, M. R., Agrawal, N., Kim, S. F., Cascio, M. B., Fujimuro, M., Ozeki, Y., Takahashi, M., Cheah, J. H., Tankou, S. K., Hester, L. D., Ferris, C. D., Hayward, S. D., Snyder, S. H., Sawa, A. (2005) S-nitrosylated GAPDH initiates apoptotic cell death by nuclear translocation following Siah1 binding. *Nat. Cell Biol.* **7**: 665-674.
- [90] Rhee, K. Y., Erdjument-Bromage, H., Tempst, P., Nathan, C. F. (2005) S-nitroso proteome of *Mycobacterium tuberculosis*: Enzymes of intermediary metabolism and antioxidant defense *Proc. Natl. Acad. Sci. U.S.A.* **102**: 467-472.

PARTE V

“Sequenziamento *de novo* delle subunità α e β della ficocianina di *Aphanizomenon flos-aquae* mediante spettrometria di massa tandem MALDI-TOF/TOF”

ABSTRACT

Aphanizomenon flos-aquae (AFA) è un'alga verde-blu (cianofita) unicellulare di acqua dolce molto ricca di ficocianina (PC), una ficobiliproteina coinvolta nel processo di fotosintesi e dalle dimostrate capacità antiossidanti ed antinfiammatorie. La PC è composta da due subunità (α e β) le cui sequenze amminoacidiche, per la specie studiata, non sono state estensivamente determinate. In questo lavoro le due catene polipeptidiche α e β sono state caratterizzate attraverso l'ausilio della spettrometria di massa con un tipico approccio "bottom-up". Un estratto di AFA è stato dapprima separato cromatograficamente grazie all'ausilio di una colonna di idrossiapatite e l'eluato è stato sia utilizzato per la determinazione della massa intatta delle sue componenti via MALDI-TOF-MS, che risolto elettroforeticamente (SDS-PAGE). Le bande ottenute dal gel sono state digerite con tripsina per la caratterizzazione dei peptidi attraverso spettrometria di massa MALDI-TOF/TOF. La strategia ha permesso di ottenere un "peptide mass fingerprinting" per ciascuna delle due subunità ed è stato realizzato un sequenziamento "de novo" dei peptidi più abbondanti attraverso dissociazione indotta da collisione (CID) a bassa energia. Al fine di ottenere un maggior numero di informazioni, sia incrementando l'efficienza della frammentazione che la semplicità di interpretazione degli spettri, i digeriti triptici sono stati, in alcuni casi, anche derivatizzati utilizzando lo SPITC (4-sulfofenil-isotiocianato). Il risultato di questi esperimenti ha portato a descrivere la sequenza amminoacidica di diversi peptidi triptici delle PC ed al raggiungimento di coperture di sequenza del 60% per le due proteine. Le ricerche effettuate nei database mostrano come le subunità α e β della PC condividano dei peptidi in comune con altre specie di cianobatteri più ampiamente caratterizzate e come sia possibile ipotizzare in *Aphanizomenon* la presenza di isoforme differenti per le PC.

CAPITOLO 1

INTRODUZIONE

1.1 *Aphanizomenon flos-acquae*

I cianobatteri (Fig. 1.1) costituiscono un ampio gruppo di organismi di notevole importanza scientifica in quanto procarioti fotosintetici e supposti precursori ancestrali dei cloroplasti presenti nelle piante superiori [1]. Il phylum

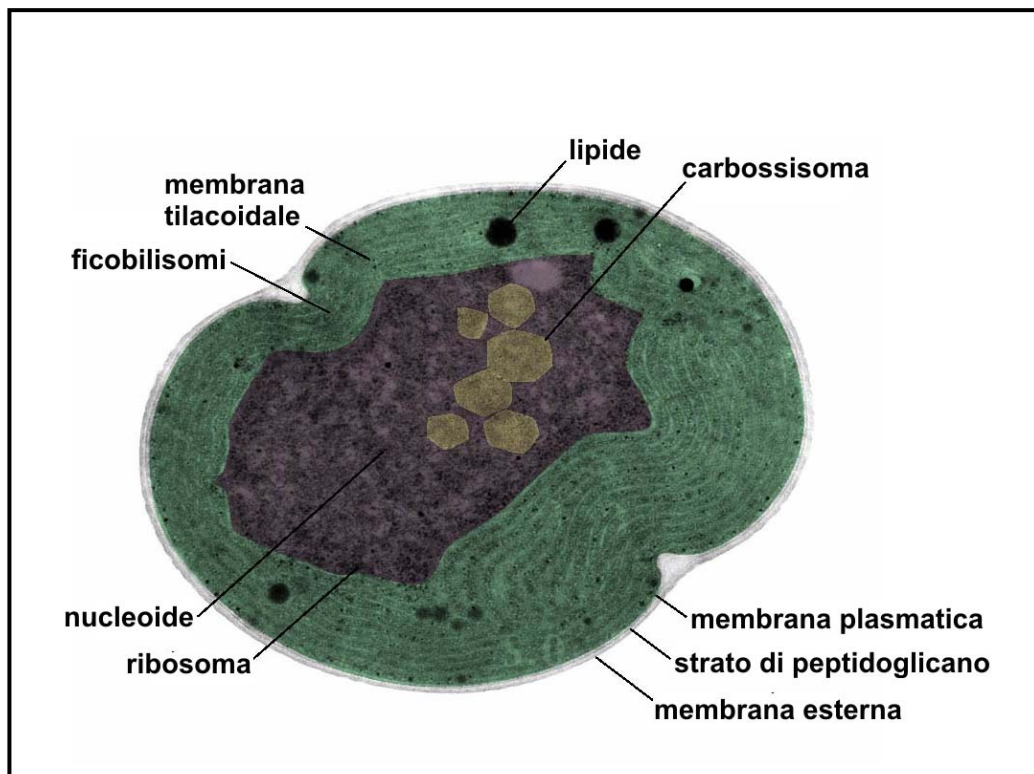


Figura 1.1 Rappresentazione schematica di una tipica cellula di cianobatterio.

comprende una grande varietà di specie differenti a partire da organismi unicellulari, come *Synechocystis*, fino ad arrivare a quelli costituiti da complesse catene filamentose multicellulari, molti dei quali hanno differenziato cellule specializzate quali le eterocisti negli azoto-fissatori *Anabaena* e *Nostoc*.

Aphanizomenon flos-aquae (AFA) è un'alga verde-blu (cianofita) unicellulare di acqua dolce che cresce spontaneamente nel lago "Upper Klamatah" (Oregon, USA). L'AFA (Fig. 1.2) viene comunemente utilizzata come cibo altamente nutriente dalle proprietà corroboranti per la salute [2, 3].



Figura 1.2 *Aphanizomenon flos-aquae*.

Seppur esistono dei lavori controversi [4, 5], diversi studi hanno documentato il valore terapeutico di questa micro-alga [6, 7] anche se il meccanismo molecolare alla base delle sue proprietà benefiche non sia ancora stato determinato. L'AFA rappresenta una ricca fonte (circa il 15% del peso secco dell'alga) della proteina fotosintetica ficocianina (PC) che è stata descritta, da evidenze sia *in vitro* che *in vivo* a carico della micro-alga *Spirulina platensis*, essere un composto naturale dalla potente azione antiossidante [8, 9] ed antinfiammatoria [10, 11]. L'azione antiossidante sembra essere legata ad un'inibizione selettiva da parte della C-ficocianina della cicloossigenasi-2 (Cox-2) coinvolta nel pathway metabolico dei mediatori chimici dell'infiammazione. La PC è una ficobiliproteina altamente fluorescente e solubile in acqua in grado di garantire ai cianobatteri la possibilità di assorbire luce a lunghezze d'onda specifiche grazie al suo legame con molecole pigmento. Nello specifico il cromoforo coniugato alla PC è indicato come ficocianobilina (PCB, Fig. 1.3) e possiede una struttura chimica simile a quella

della bilirubina che gli consente come quest'ultima di funzionare come "scavenger" per le specie reattive dell'ossigeno (ROS) [12, 13].

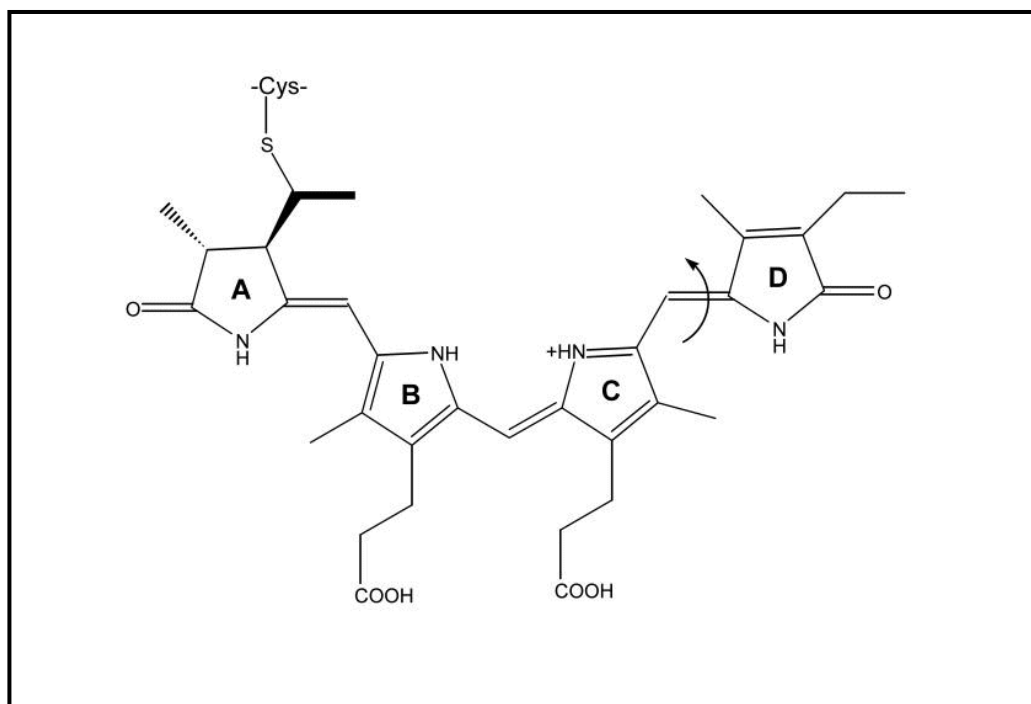


Figura 1.3 Struttura chimica del pigmento ficocianobilina.

In questo contesto, un recente studio ha descritto le proprietà antiossidanti degli estratti naturali di AFA nel proteggere gli eritrociti umani ed i campioni di plasma contro il danno ossidativo. [14]. La ficocianina trova inoltre impiego come additivo colorante in diversi cibi [15], nell'industria cosmetica [16] e nella ricerca biomedica [17].

1.2 Sistema antenna nei cianobatteri: i Ficobilisomi

I **ficobilisomi** (PBSs) sono grandi complessi multiproteici che hanno masse molecolari di circa $5\text{-}20 \times 10^6$ Da (a seconda dell'organismo e delle condizioni di crescita) e sono generalmente posizionati sul lato stromale (protoplasmatrico) delle membrane tilacoidali dei cianobatteri (Fig. 1.1) [18]. Il loro compito principale è quello di catturare l'energia luminosa e convogliarla al centro di reazione dei fotosistemi I e II (PSI, PSII) per poter essere utilizzata nella catena di trasporto elettronico che, in ultima analisi, ne consente la sua trasformazione in energia chimica. Micrografie elettroniche ottenute in seguito a congelamento-frattura mostrano righe di PBSs emidiscoidali con una periodicità che riflette quella delle

particelle dimeriche del PSII, che si osservano sul lato citoplasmatico della membrana tilacoidale. Questo suggerisce che il supercomplesso PBS-PSII è composto da due complessi PSII e da un PBS, riarrangiamento confermato anche da analisi biochimiche e studi di trasferimento di energia. L'organizzazione del supercomplesso PBS-PSII in righe strettamente ordinate non permette solo una alta densità di impaccamento dei PBSs, ma anche un trasferimento ad alta efficienza dell'energia di eccitazione tra le unità adiacenti di PSII. Questo sistema di fibre conducenti energia consente una più efficiente distribuzione dell'energia lungo tutta la superficie delle membrane tilacoidali, ed aiuta ad ottimizzare la raccolta della luce. Tuttavia l'associazione del PBS con il PSII dimero sembra essere dinamica, infatti sono stati osservati complessi PBSs distribuiti più o meno a caso sulla superficie delle membrane tilacoidali dei cianobatteri; l'organizzazione dei PBSs in righe ordinate sembra dipendere dalle condizioni luminose ricevute durante la crescita. La dissociazione totale o parziale del supercomplesso PBS-PSII avviene tramite una ridistribuzione laterale dei PBSs ("stato di transizione"), che a sua volta altera la distribuzione dell'energia tra i fotosistemi. Analogamente a quanto avviene nelle piante, mediante questo stato di transizione, i cianobatteri rispondono ad una preferenziale eccitazione del PSII o del PSI potenziando il trasferimento dell'energia di eccitazione al fotosistema meno attivo. I ficobilisomi sono costituiti da due elementi strutturali: un **core centrale** e dei **rods** periferici che sono disposti in forma emidiscoidale intorno al core, che nella maggior parte dei PBSs è costituito da tre cilindri (Fig. 1.4). I rods periferici si emanano radialmente dalle superfici laterali del core che non sono in contatto con la membrana tilacoidale. I polipeptidi costituenti i PBSs possono essere raggruppati in tre classi: 1) *ficobiliproteine* (80% in massa); 2) *polipeptidi linker*; 3) *proteine associate ai ficobilisomi* (ad esempio FNR = NADP⁺ ossidoriduttasi). La composizione proteica dei PBSs può variare notevolmente tra i diversi ceppi di cianobatteri [18]. Sulla base del loro caratteristico assorbimento, le **ficobiliproteine** sono state suddivise in quattro classi: **ficoeritrocianine** (PEC) con $\lambda_{\max} = 575$ nm, **ficoeritrine** (PE) $\lambda_{\max} = 565-575$ nm, **ficocianine** (PC) $\lambda_{\max} = 615-640$ nm, **alloficocianine** (APC) $\lambda_{\max} = 650-655$ nm. Tali differenze spettroscopiche possono essere dovute a tre fattori: a) legame con cromofori diversi, b) diversi tipi di legame chimico con lo stesso cromoforo, c) diverso intorno del cromoforo. Queste differenze spettroscopiche corrispondono anche alla diversa localizzazione strutturale di tali complessi, e quindi alla loro posizione nell'ambito

del flusso di energia dalla periferia al core centrale: infatti, le ficobiliproteine che assorbono a λ maggiori (PC) si trovano nella zona periferica dei rods, PEC e PE in prossimità del core, mentre APC, che è il complesso dei ficobilisomi con il max di assorbimento più spostato verso il rosso, è il maggior componente del core, ricevendo l'energia dai rods periferici e trasferendola alle proteine che legano clorofilla della membrana tilacoidale.

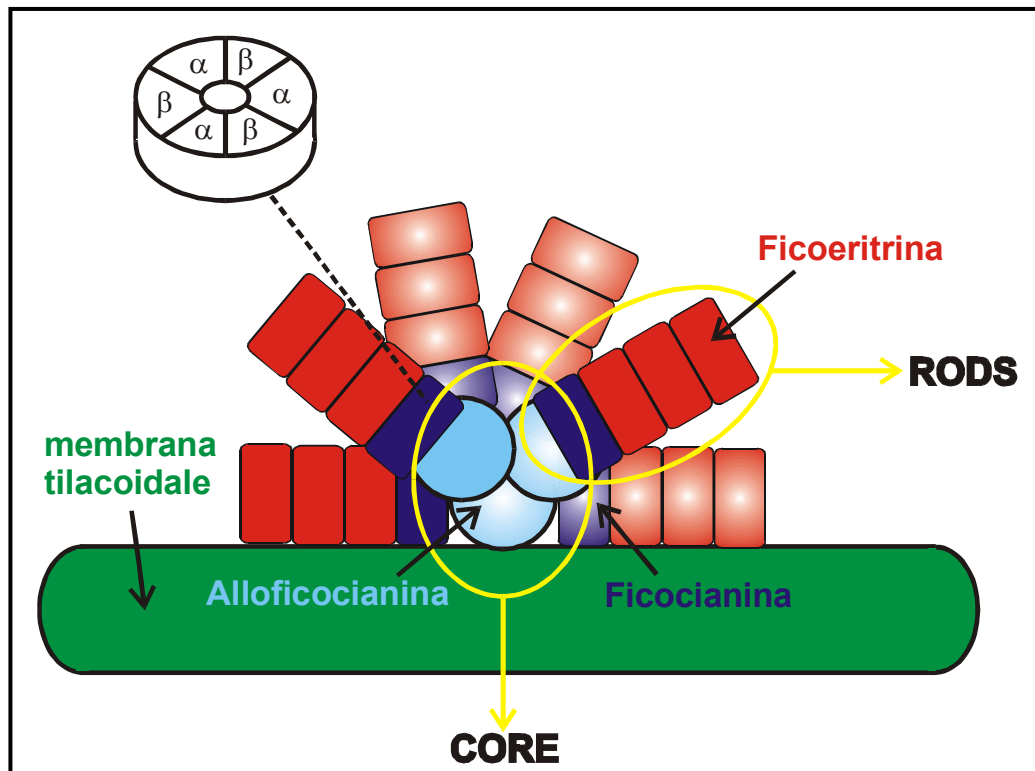


Figura 1.4 Rappresentazione schematica del Ficybilisoma (PBS), il complesso antenna dei cianobatteri.

Per quanto riguarda la struttura fine, le ficobiliproteine sono etero-monomeri costituiti da due differenti subunità, α e β , presenti in stechiometria equimolare ($\alpha\beta$). Tali subunità differiscono in massa molecolare, sequenza amminoacidica e contenuto dei cromofori. L'unità fondamentale per tutti i PBSs è un trimero stabile ($\alpha\beta$)₃ che forma un aggregato di forma toroidale con un diametro di 11 nm e uno spessore di 3-3.5 nm, con una cavità centrale di 3 nm (inserto di Fig. 1.4). Tali dischi trimerici si aggregano faccia a faccia a formare degli esameri, con o senza l'inclusione nella cavità di un peptide linker. Questi ultimi costituiscono circa il 20% del contenuto proteico totale dei PBSs e sono responsabili della stabilizzazione

degli esameri di PE, PC e PEC e, a volte, dell'unione di unità esameriche a costituire aggregati di dimensioni maggiori come i rods o i cilindri del core. Ad esempio, le alloficocianine assemblano il core del ficobilisoma con l'assistenza di tre tipi di linkers: la grande proteina L_{CM} , il linker L_{RC} , che connette il core con i rods periferici e i piccoli linkers L_C . Una nota a parte merita L_{CM} o ApcE, dal nome del corrispondente gene, in quanto è la più grande cromoproteina dei PBSs con i suoi 70 -128 kDa (a seconda degli organismi) di peso molecolare. Si ritiene che L_{CM} ancori i PBSs alla membrana tilacoidale, di qui il nome di peptide-ancora, quindi esso gioca un ruolo centrale nell'assemblaggio e nell'architettura dei PBSs, e nel trasferimento dell'energia dal ficobilisoma alla clorofilla *a* associata con le antenne interne del PSII (CP43 e CP47).

Le biliproteine dei cianobatteri possono essere ottenute come prodotti di dissociazione dei ficobilisomi. Quando le cellule procariotiche sono rotte ed il contenuto cellulare si trova in un mezzo acquoso a bassa forza ionica, i ficobilisomi dissociano nei vari componenti, ed è possibile ottenere ed analizzare le biliproteine, con o senza i linkers associati. La relativa stabilità dei complessi biliproteina-linker varia tra le biliproteine e le diverse fonti. L'alloficocianina, con un massimo a 650 nm, si ritrova a pH neutro come un trimero costituito da tre polipeptidi α e tre polipeptidi β , ognuno dei quali lega un cromoforo (bilina). Tre monomeri $\alpha\beta$ si assemblano in forma di anello a costituire un trimero $\alpha_3\beta_3$ in una *threefold symmetry*. La C-ficocianina si ritrova come soluzione di $\alpha_3\beta_3$, $\alpha_6\beta_6$ ed altri oligomeri. L'esamero $\alpha_6\beta_6$ ha forma discoidale, ed è costituito da due trimeri associati faccia a faccia. I rods sono costituiti dall'associazione faccia a faccia di questi dischi. Il polipeptide α lega una bilina, mentre il polipeptide β due. La C-ficoeritrina, che anch'essa si associa in dischi, è costituita da un polipeptide α con due cromofori ed un polipeptide β con tre cromofori. Il peso molecolare di questi polipeptidi α e β varia in un range di massa molecolare compreso tra 15.000 e 20.000 Da.

I monomeri ($\alpha\beta$) dell'alloficocianina e della C-ficocianina sono interessanti per lo studio delle biliproteine perché rappresentano la minima unità avente tutte le biline. Si possono ottenere monomeri omogenei e stabili di queste biliproteine in condizioni di pH leggermente acido o in presenza di anioni caotropici quali tiocianato o perclorato [19, 20]. Le biliproteine possono essere completamente dissociate nei singoli polipeptidi α e β usando agenti denaturanti [21]. La determinazione della sequenza amminoacidica di varie biliproteine è stato un

passo importante verso la comprensione delle loro proprietà. Da tali studi è risultata una significativa omologia di sequenza tra le varie biliproteine ed inoltre la posizione dei cromofori è altamente conservata. Le **biline** sono dei tetrapirroli a catena aperta legati covalentemente all'apoproteina mediante legami tioeteri su particolari residui di cisteina. Nei cianobatteri sono conosciuti sia strutturalmente che spettroscopicamente nove differenti **bili-pigmenti**, che differiscono nel numero e nell'organizzazione dei doppi legami coniugati. I quattro pigmenti principali (presenti anche nelle alghe rosse) sono: la *ficocianobilina* di colore blu (PCB); la *ficoeritrobilina* di colore rosso (PEB); la *ficourobilina* di colore giallo (PUB); la *ficobiliviolina* di colore porpora (PXB; chiamata anche criptoviolina). Le diverse proprietà di assorbimento delle ficobiline sono causate principalmente dalle differenze nel numero di doppi legami coniugati e dalle catene laterali nei gruppi

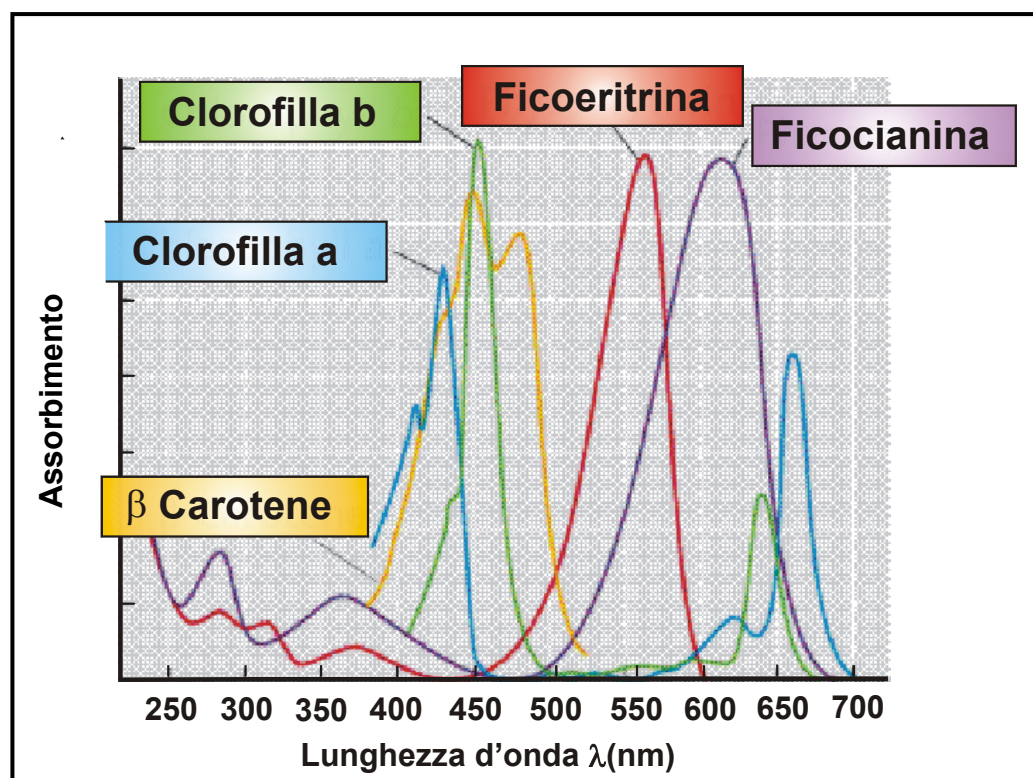


Figura 1.5 Spettri di assorbimento dei principali pigmenti fotosintetici.

prostetici tetrapirrolici (Fig. 1.5). La C-ficocianina e l'alloficocianina legano ficocianobiline, mentre la C-ficocianoeritrina lega le ficoeritrobiline. Nessuno ione metallico è associato alle biline. Quando i linkers dei rods legano biline, sono chiamati polipeptidi γ . I linkers che fanno parte dei rods della C-ficocianina e della C-ficoeritrina sono probabilmente incolori.

Nel processo fotosintetico le biliproteine hanno la funzione di catturare l'energia luminosa in regioni dove l'assorbimento della clorofilla *a* è scarso, conseguentemente questa energia migra con grande efficienza dal punto di assorbimento verso il centro di reazione. È importante notare che le biliproteine assorbono i fotoni su un ampio range di energie. Il selettivo assorbimento della luce delle varie biline si ottiene inizialmente grazie a differenze nel loro sistema di doppi legami coniugati. La ficocianobilina ha un grado di coniugazione maggiore della ficoeritrobilina, cosicché soluzioni di ficocianina vicine a pH neutro assorbono ad energia minore (massimo a 610-620 nm) delle ficoeritrine (massimo a 545-565 nm). Accanto a biliproteine che legano soltanto una particolare bilina per la cattura della luce, esistono biliproteine che possiedono due differenti biline, con conseguente drammatico aumento nel range dell'assorbimento dei fotoni. Nessuna biliproteina batterica finora studiata presenta tre diverse biline, sebbene questo è stato scoperto in *Cryptomonas*. A partire dalla struttura chimica delle biline, ulteriori fattori vanno a modulare le caratteristiche di cattura della luce e di migrazione dell'energia influenzando lo spettro delle biline stesse. Ad esempio, l'apoproteina mantiene le biline in una conformazione aperta che massimizza l'assorbimento nella regione visibile dello spettro; al contrario, senza l'influenza dell'apoproteina, le biline assumerebbero una conformazione più ciclica e avrebbero un basso assorbimento nel visibile.

CAPITOLO 2

MATERIALI E METODI

2.1 Purificazione delle PC

Le PC sono state purificate cromatograficamente a partire da un estratto secco di AFA gentilmente concesso dalla Nutratec (Urbino). 500 mg di estratto sono stati risospesi utilizzando 50 ml di un tampone 100 mM Na-fosfato a pH 7, portato poi al 50% di saturazione con solfato d'ammonio e lasciato incubare per 1 h a 4 °C. Il precipitato ottenuto è stato recuperato per centrifugazione a 10.000 g per 30 min. a 4 °C; il surnatante privo di colore è stato eliminato ed il precipitato blu è stato dissolto in un piccolo volume di tampone 5 mM Na-fosfato a pH 7 e dializzato overnight a 4 °C contro lo stesso tampone. L'estratto di PC dializzato è stato infine caricato su di una colonna di idrossiapatite di 2,5 x 25 cm (Bio-Rad Laboratories, CA, USA) e diverse frazioni sono state ricavate eluendo con tampone Na-fosfato a pH 7 a crescente forza ionica (da 5 a 150 mM). Le frazioni che hanno mostrato un rapporto degli assorbimento 620/280 nm maggiore di quattro sono state prelevate e portate a saturazione con solfato d'ammonio. Dopo centrifugazione, il precipitato blu è stato recuperato e dissolto in 150 mM Na-fosfato a pH 7 e dializzato nuovamente contro lo stesso campione. Le PC così purificate sono state conservate al buio a 4 °C.

2.2 SDS-PAGE

Per separare le due subunità α e β della PC prima dell'analisi proteomica si è realizzata un'elettroforesi monodimensionale. L'estratto puro di PC arricchito dalla separazione cromatografica è stato disciolto in 2% SDS ed 1% β -mercaptoetanololo e fatto bollire per 5 min. L'analisi è stata condotta utilizzando un Laemli gel [22] al 15% w/v di acrilammide (AA) evidenziando il valore dei pesi molecolari per confronto con un marker di proteine di pesi molecolari compresi tra i 14.400 Da e 200.000 Da. Il gel è stato fatto correre ad una corrente costante di 12 mA a temperatura ambiente ed infine colorato con Blue Coomassie colloidale (Bio-Rad, Laboratories, CA, USA).

2.3 Digestione triptica delle proteine in-gel

Le bande colorate con Blue Coomassie sono state ritagliate dal gel usando un sottile bisturi e soggette a digestione triptica in accordo con il metodo proposto

da Shevchenko et al. [23]. I pezzi di gel, posti a 4° C, sono stati lasciati imbibire nel tampone di digestione contenente NH_4HCO_3 50 mM e 12,5 ng/ μl di tripsina (modified porcine trypsin, sequencing grade; Promega, Madison, WI, USA). Dopo 30 minuti l'eccesso di liquido è stato rimosso e sono stati aggiunti 20 μl di NH_4HCO_3 50 mM lasciando i frammenti di gel ad incubare overnight ad una temperatura costante di 37 °C. Al termine dell'incubazione è stato recuperato il sopranatante contenente i peptidi triptici; l'estratto, portato a secco centrifugando sotto vuoto, è stato conservato a -20 °C fino al momento dell'uso. Prima dell'utilizzo delle miscele peptidiche per le analisi di spettrometria di massa queste sono state disciolte in 10 μl FA 5% (acido formico).

2.4 Determinazione della massa intatta della proteina: MALDI-TOF-MS

L'estratto di PC è stato ulteriormente purificato, prima dell'analisi MS, utilizzando una nano-colonna impaccata con la resina Poros R1 (PerSeptive Biosystems, Framingham, MA, USA) in un puntale GELoader (Eppendorf, Amburgo, Germania) preventivamente schiacciato all'estremità. 1 μl di soluzione PC è stato utilizzato per l'analisi ed eluito direttamente dalla colonna usando la soluzione contenente la matrice, nello specifico: acido sinapinico (SA) 20 $\mu\text{g}/\mu\text{l}$ in acetonitrile (ACN), 0.1% acido trifluoroacetico (TFA) (70:30 v/v). Per la deposizione del campione sul target si è utilizzato il metodo "sandwich": si sono depositati sul target 0.5 μl di una soluzione SA 20 $\mu\text{g}/\mu\text{l}$ in acetone i quali evaporando istantaneamente hanno formato una patina sulla quale successivamente si sono applicati 0,5 μl 2% TFA e l'eluato della micropurificazione contenente le proteine da analizzare disciolte nella matrice. Gli spettri di massa sono stati acquisiti utilizzando lo strumento "Voyager-DE STR BioSpectrometry Workstation" (Applied Biosystem, Foster City, CA). Gli ioni sono stati analizzati in modalità positiva utilizzando la "delayed extraction". Lo strumento è stato calibrato esternamente utilizzando lo standard "Protein Calibration Standard I" della Bruker-Daltonics (Brema, Germania) contenente una miscela di prodotti proteici con pesi compresi tra i 3.000 Da ed i 20.000 Da. Tutti gli spettri acquisiti sono stati analizzati con l'ausilio dei software MoverZ (Proteometrics, Winnipeg, Canada) e Data Explorer versione 4.5 (Applied Biosystem, Foster City, CA).

2.5 Spettrometria di massa MALDI-TOF/TOF: Peptide Mass Fingerprinting (PMF) e De Novo sequencing

Il metodo descritto da Gobom et al. [24] è stato utilizzato per eliminare i contaminanti salini e concentrare i peptidi triptici. Usando dei puntali GELoader (Eppendorf, Amburgo, Germania) preventivamente schiacciati all'estremità si sono impaccate delle nanocolonne dal diametro di 5mm con la resina Poros R2 (PerSeptive Biosystems, Framingham, MA, USA). Attraverso l'ausilio di una siringa che ha permesso di applicare una lieve pressione le soluzioni liquide sono state sospinte attraverso le colonne. Le colonne sono state dapprima pulite con 20 µl 100% ACN, equilibrate con 20 µl di acido formico (FA) al 5%, poi caricate con le miscele peptidiche ed infine lavate con 20µl FA 5%. L'eluizione dei peptidi dalla colonna è stata realizzata direttamente sul target MALDI utilizzando 0.5 µl di soluzione CHCA (acido 3-(3-ciano-4-idrossicinnamico 20µg/µl in ACN/0.1% TFA v/v, 70:30 v/v). I campioni sono stati analizzati usando lo strumento "Applied Biosystems 4700 Proteomics Analyzer MALDI-TOF/TOF" (Applied Biosystems, Framingham, MA, USA). Gli ioni positivi sono stati esaminati in modalità reflectron ed il gas utilizzato per il CID nella cella di collisione è stato aria atmosferica. L'analisi dei risultati è stata fatta utilizzando il software Data Explorer versione 4.5 (Applied Biosystem, Foster City, CA). La calibrazione degli spettri è stata realizzata internamente tramite l'utilizzo dei segnali relativi ai prodotti dell'autodigestione della tripsina (m/z 842,51 e 2.211,10). Quando questo non è stato possibile si è ricorsi ad una calibrazione esterna con "Peptide Calibration Standard II" Bruker-Daltonics (Brema, Germania).

Per l'identificazione delle proteine è stato utilizzato il software MASCOT (<http://www.matrixscience.com>) che ha permesso di eseguire le ricerche in un database non-ridondante contenente sequenze proteiche (NCBI nr database). I seguenti parametri sono stati utilizzati per compiere le ricerche nel database: accuratezza nella determinazione della massa monoisotopica del peptide pari a 70 ppm, tolleranza sulla massa dei frammenti MS/MS $\pm 0,6$ Da, numero dei possibili siti di taglio della tripsina non utilizzati pari ad uno, completa carbamidometilazione delle cisteine e parziale ossidazione delle metionine. Tutte le sequenze amminoacidiche ottenute, sia per i peptidi derivatizzati e non, sono state controllate manualmente.

2.6 Derivatizzazione chimica con 4-sulfofenil-isotiocianato (SPITC)

Lo SPITC è stato disciolto in una soluzione 20 mM NaHCO_3 pH 8,6 ad una concentrazione di 10 $\mu\text{g}/\mu\text{l}$. Poche pmoli della miscela peptidica da derivatizzare sono state addizionate con 7 μl della soluzione SPITC e lasciate incubare per 30 min. a 50 °C. Al termine dell'incubazione la reazione è stata bloccata aggiungendo 1 μl 5% TFA. I campioni così ottenuti sono stati micropurificati prima dell'analisi in MS come descritto in sezione 2.5.

CAPITOLO 3

RISULTATI

L'estratto cellulare di AFA, dopo purificazione cromatografica utilizzando una colonna di idrossipatite, è stato utilizzato per la determinazione della massa molecolare intatta delle proteine in esso contenute.

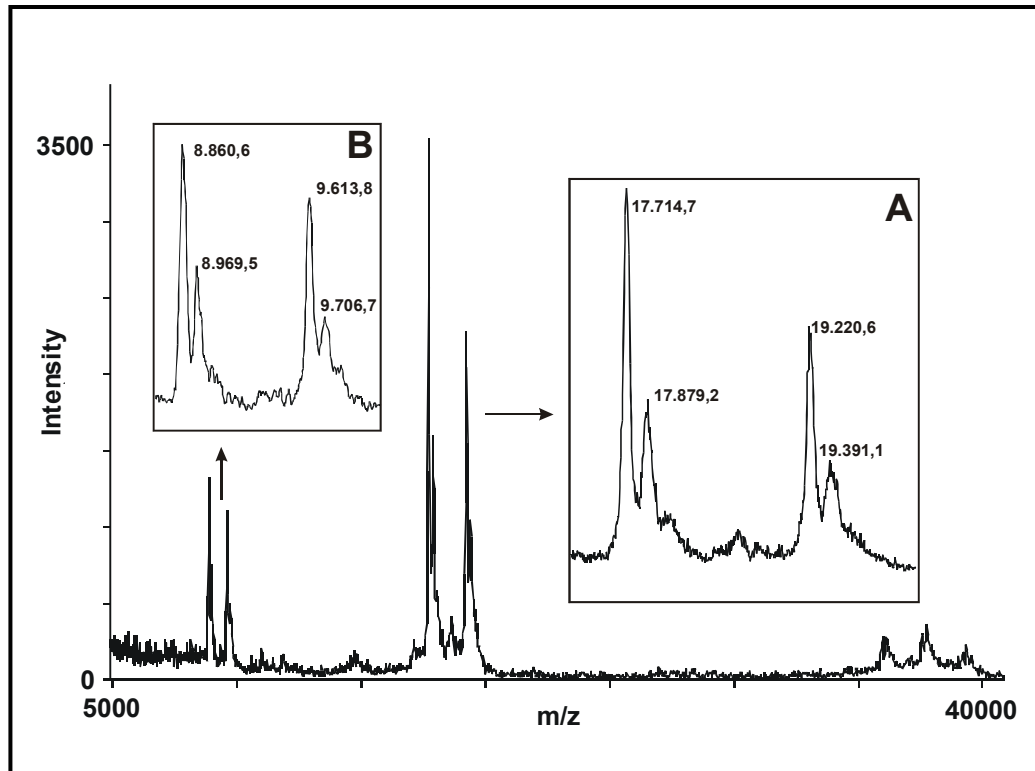


Figura 3.1 Spetto di massa MALDI-TOF dell'estratto di PC dopo purificazione cromatografica su colonna di idrossipatite.

In Fig. 3.1 è mostrato lo spettro di massa MALDI-TOF ottenuto dall'analisi; i due picchi di maggiore intensità (17.714,7 Da e 19.220,6 Da, inserto A in Fig. 3.1) possono essere attribuiti alle catene α e β della ficocianina comparando il valore dei pesi molecolari ottenuti con quelli dedotti dalle sequenze di DNA del cianobatterio *Nostoc sp. PCC 7120* (Tabella 3.1). E' interessante notare come per entrambi i due picchi principali siano presenti due picchi di spalla che differiscono per meno di duecento unità di massa (17.879,2 Da e 19.391,1 Da, inserto A in Fig. 3.1). Nella regione a più bassi pesi molecolari sono anche presenti i picchi corrispondenti agli ioni bicarica delle ficocianine stesse (8.860,6 Da e 9.613,8 Da, inserto B in Fig. 3.1).

Peso molecolare osservato	Peso molecolare atteso	Proteina [organismo omologo]	NCBI Accession number
17.713,7	17.456,5	cpcA [Nostoc sp. PCC 7120]	gi 9957319
19.219,6	18.327,8	cpcB [Nostoc sp. PCC 7120]	gi 38894

Tabella 3.1 Identificazioni delle PC mediante misura del peso molecolare

L'estratto purificato dopo valutazione della massa intatta delle proteine è stato risolto elettroforeticamente utilizzando una SDS-PAGE monodimensionale (Fig. 3.2) che ha evidenziato la presenza di due distinte bande di peso molecolare apparente pari a 18.500 Da (subunità α) e 21.900 Da (subunità β).

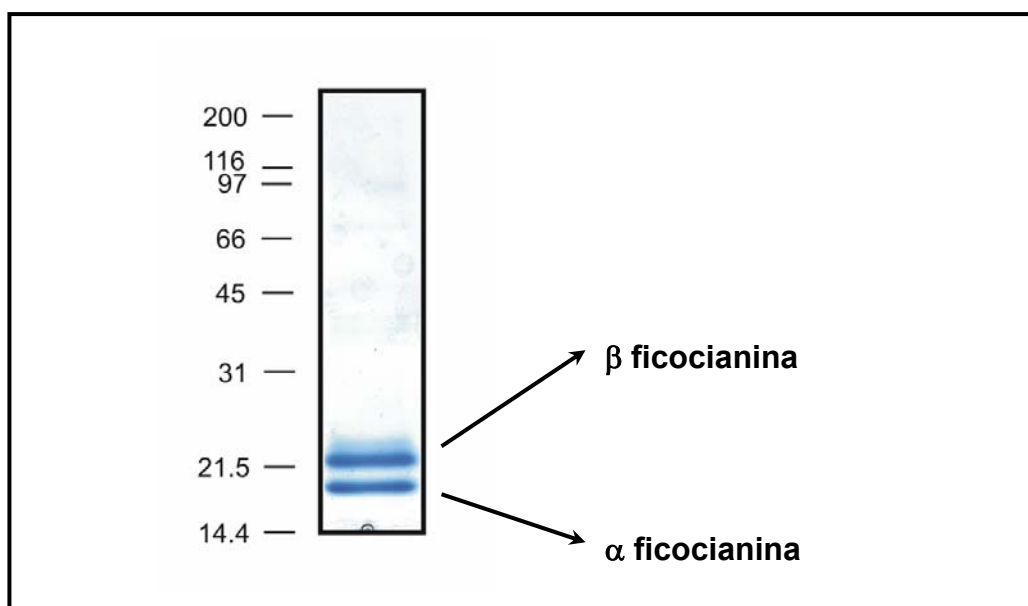


Figura 3.2 Separazione elettroforetica SDS-PAGE dell'estratto purificato di PC.

Le bande in esame sono state escisse, sottoposte a digestione triptica in-gel ed i peptidi ottenuti analizzati attraverso uno spettrometro di massa MALDI-TOF/TOF. Sebbene solo poche informazioni di sequenza relative ad *Aphanizomenon flos-aquae* siano presenti nei database, ed in particolare solo due brevi tratti di sequenza per le catene α e β della ficocianina (Tabella 3.2), si è valutato potesse essere interessante confrontare il pattern peptidico ottenuto con quello generato dalla digestione “in silico” di proteine dello stesso tipo, appartenenti ad altre specie, presenti nelle banche dati.

Proteina [<i>Aphanizomenon flos-aquae</i>]	NCBI Accession number	Sequenza
Ficocianina subunità α	gi 11121363	MKTPITEAIASADTQG AGDASVLDDRCLNGLRETY
Ficocianina subunità β	gi 11121362	QALGTPGSSVAVGVGKMKE AAIAIVNDPNGITKGDCSSLV SELASYFDRAAAAVV

Tabella 3.2 Informazioni di sequenza relative alle due subunità della ficocianina di AFA presenti in banca dati.

L'interrogazione dei database con i soli dati relativi al PMF non hanno fornito nessun risultato e si è quindi proceduto ad analizzare attraverso tandem MS i picchi principali presenti nello spettro di massa acquisito. L'utilizzo del software Mascot per l'analisi combinata dei dati PMF delle proteine oggetto di studio e delle informazioni di sequenza ottenute tramite MS/MS ha portato all'identificazione positiva dell' α e della β PC attraverso il riconoscimento statisticamente significativo delle proteine in organismi affini (Tabella 3.3).

Proteina [organismo omologo]	NCBI Accession number	Mascot score	Peptidi Identificati (M+H) ⁺
Ficocianina subunità α [<i>Nostoc sp. PCC 7120</i>]	gi 9957319	250	859,56 AAASLEAAR 2.578,25 ANHGLSGQAANEANTYIDYAIN ALS
Ficocianina subunità β [<i>Anabaena mendotae</i>]	gi 51105738	271	1.791,77 GDCSSLVSEVASYFDR 1.799,87 YVTYAAIAGDASVLDDR

Tabella 3.3 Risultati dell'interrogazione delle banche dati utilizzando i dati PMF e le informazioni di sequenza ottenute con MS/MS.

Allo scopo di validare la bontà dell'identificazione fornita dal software e l'avvenuto corretto sequenziamento, si è proceduto ad un'attenta analisi manuale degli spettri MS/MS constatando l'esistenza per i peptidi identificati di un numero sufficiente di ioni diagnostici da permettere la loro completa copertura di sequenza (Fig. 3.3). Le analisi MS/MS sono state compiute selezionando gli ioni che presentavano un

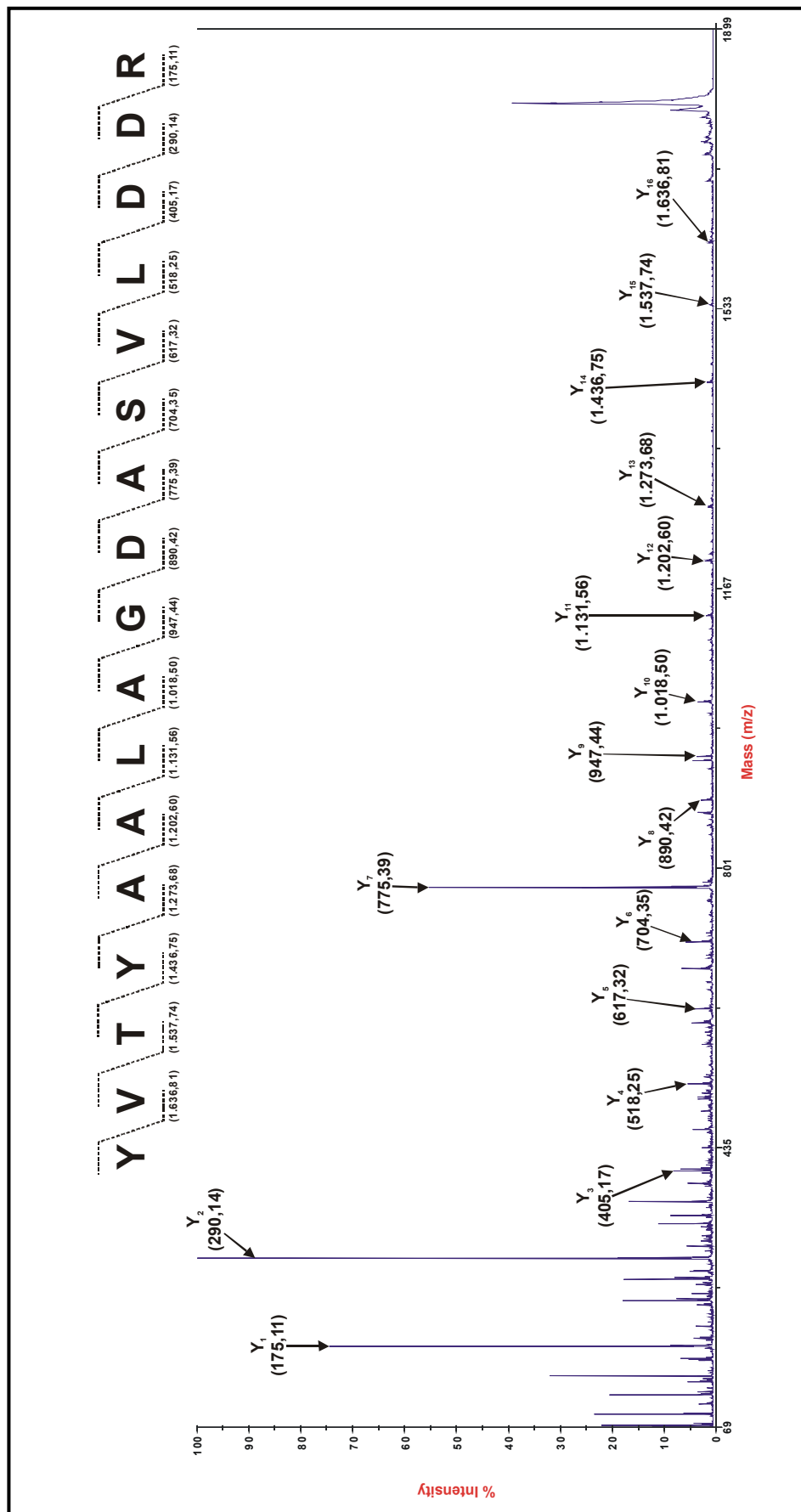


Figura 3.3 Spettro MS/MS relativo al peptide 1.799,85 appartenente alla β ficocianina di AFA.

segnale migliore nello spettro di massa. Questo non ha garantito che il risultato della frammentazione fosse egualmente buono per tutti i peptidi analizzati, ragione per cui non è stato possibile elaborare una sequenza per ognuno dei prodotti sottoposti a CID. Alla luce di tutto questo, con l'intento di riuscire a caratterizzarle in maniera più estensiva le due PC, si è pensato di utilizzare una tecnica di derivatizzazione chimica per i peptidi che potesse migliorare la qualità degli spettri MS/MS. I digeriti triptici sono stati derivatizzati utilizzando lo SPITC ed il corretto funzionamento del protocollo è stato vagliato comparando gli spettri MS delle miscele peptidiche prima e dopo la derivatizzazione (Fig. 3.4).

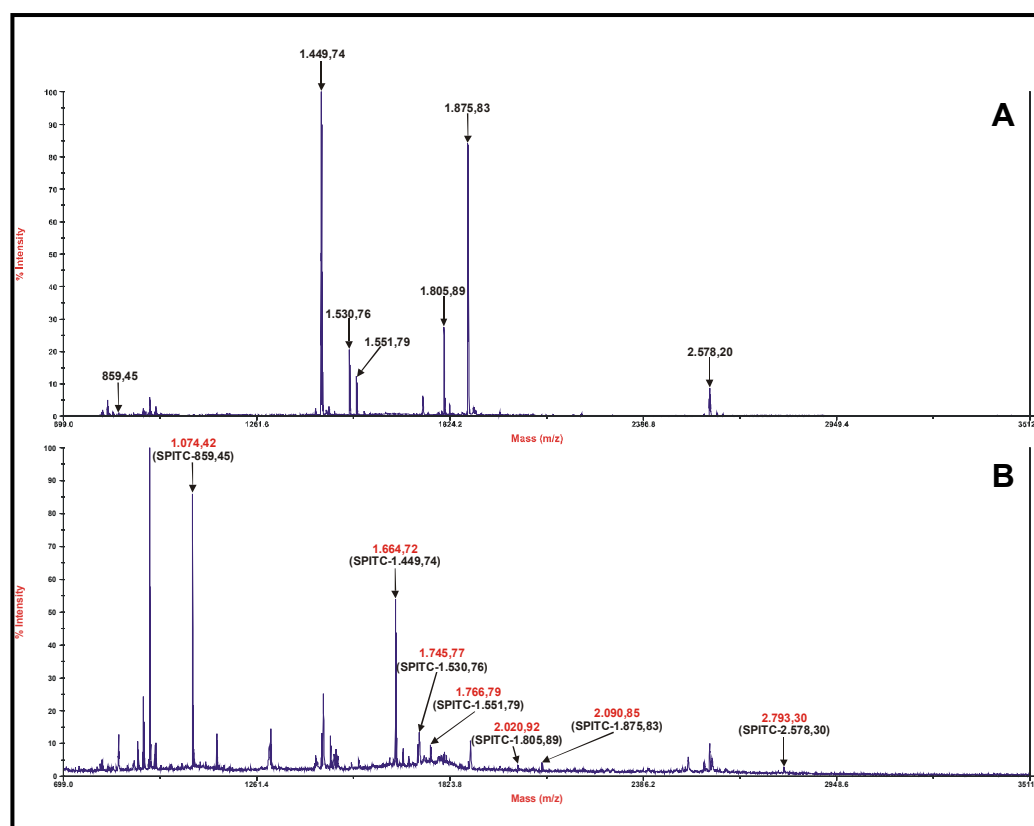


Figura 3.4 Spetto di massa MALDI-TOF/TOF del digerito triptico della ficocianina α prima (A) e dopo (B) derivatizzazione con lo SPITC.

Come si può notare l'aggiunta dello SPITC ha portato ad uno shift dei valori di massa pari a 214,97 Da che ha permesso di riconoscere i peptidi che sono stati derivatizzati i quali sono stati poi frammentati e sequenziati. La qualità degli spettri che si è ottenuta utilizzando lo SPITC è di gran lunga superiore se confrontata con quella degli spettri che si sono ottenuti senza derivatizzazione (Fig. 3.5). L'utilizzo di questa strategia ha permesso di confermare il corretto sequenziamento dei peptidi mostrati in Tabella 3.3 e di ottenere le sequenze amminoacidiche anche

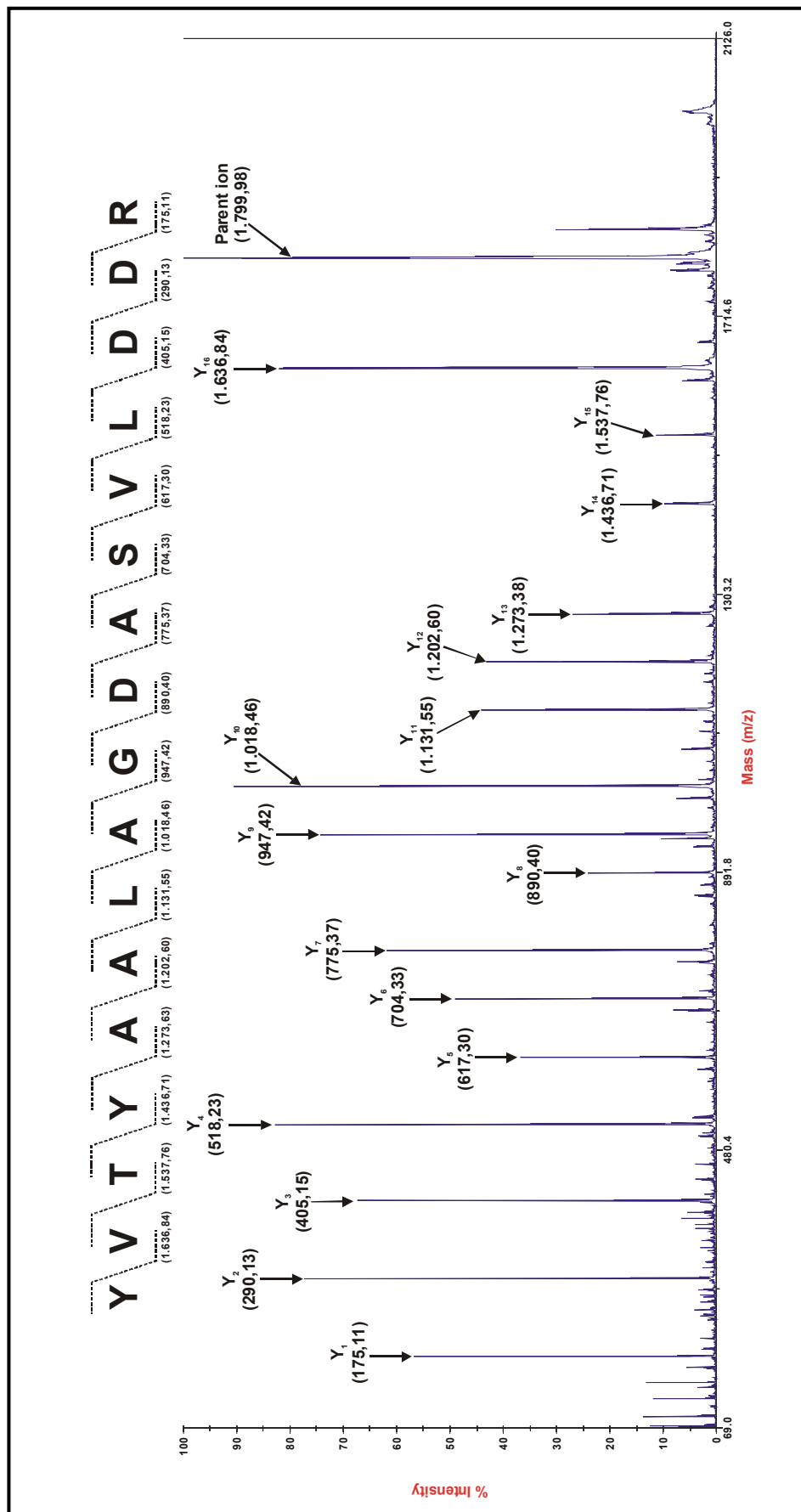


Figura 3.5 Spettro MS/MS relativo al peptide 2.014,84 (SPITC-1.799,85) appartenente alla β ficocianina di AFA.

dei peptidi che avevano prodotto in precedenza spettri di bassa qualità. Inoltre, tale strategia ha consentito anche di poter identificare e sequenziare ulteriori peptidi che non risultavano presenti nello spettro di massa originale prima della derivatizzazione. Le Tabelle 3.4 e 3.5 riassumono i dati di sequenza ottenuti in seguito all'utilizzo dello SPITC.

Proteina [<i>Aphanizomenon</i> <i>flos-aquae</i>]	MW peptide; MW SPITC-peptide	Sequenza
Ficocianina subunità α	859,45	DVGHYLR
	1.074,42	
	859,45	AAASLEAAR
	1.074,51	
	1.449,74	FLSNTELQAVDGR
	1.664,72	
	1.530,76	TPITEAIASADTQGR
	1.745,77	
	1.551,79	AQKLIDGATSAVYSK
	1.766,79	
	1.805,89	MKTPITEAIASADTQGR
	2.020,92	
	1.875,83	FPYTTSTPGNQYASDAR
	2.090,85	
	2.578,20	ANHGLSGQAANEANTYIDYAINALS
	2.793,30	

Tabella 3.4 Peptidi dell' α ficocianina sequenziati de novo dopo derivatizzazione con lo SPITC.

In Fig. 3.6 e Fig. 3.7 sono mostrati gli spettri di massa relativi a due dei peptidi sequenziati della β ficocianina. Il primo spettro, relativo al peptide 1.791,74 Da (2.006,74 Da quando derivatizzato con lo SPITC), è risultato nella sequenza GDCSSLVSEVASFYFDR omologa alla sequenza di molte altre ficocianine di tipo β presenti in diversi organismi. Nello spettro in realtà, come si può notare, esiste

Proteina [<i>Aphanizomenon</i> <i>flos-aquae</i>]	MW peptide; MW SPITC-peptide	Sequenza
Ficocianina subunità β	-	RLDVVNR
	1.086,48 *	
	-	DMEILR
	1.104,41 *	
	-	DMoxEILR
	1.120,45 *	
	1.388,75	ITSNASAIVTNAAR
	1.603,72	
	1.636,82 ‡	VTYAAIAGDASVLDDR
	-	
	1.700,76	GDRSLVSEVASYFDR
	1.915,74	
	1.734,75 ‡	GDCSSLVSEVASYFD
	-	
	1.791,74	GDCSSLVSEVASYFDR
	2.006,74	
	1.799,85	YVTYAAIAGDASVLDDR
	2.014,84	
	2.123,09	GEFLSTEQLDALTAVV??R
	2.338,05	
	2.205,17	SLFEEQPNLIAPGGQAYTNR
	2.420,02	
	2.361,17	SLFEEQPNLIAPGGQAYTNRR
	2.576,15	

Tabella 3.5 Peptidi dell' α ficocianina sequenziati de novo dopo derivatizzazione con lo SPITC. * Peptidi rilevabili solo dopo derivatizzazione. ‡ Peptidi dovuti a ulteriore frammentazione di peptidi triptici.

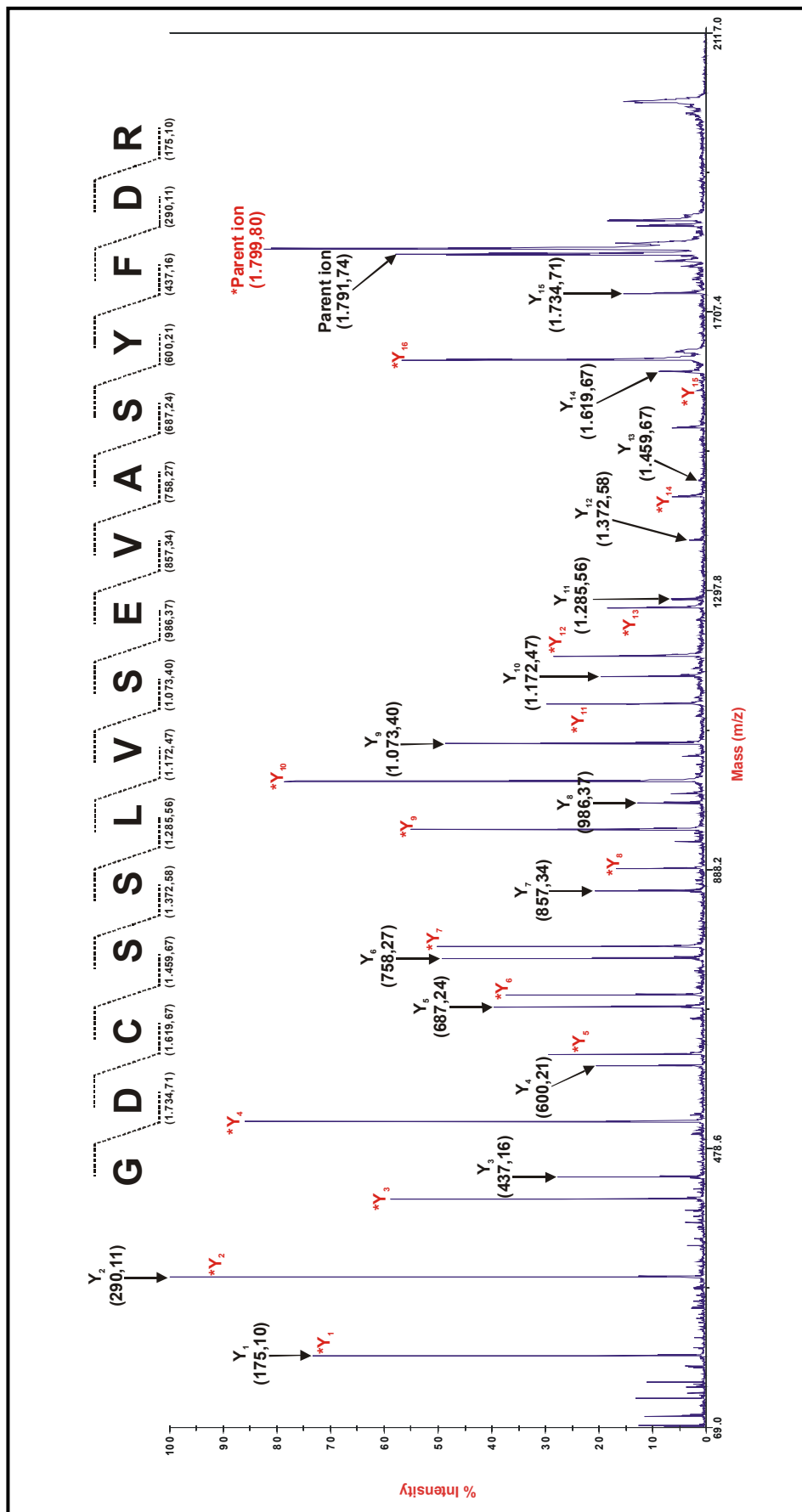


Figura 3.6 Spettro MS/MS relativo al peptide 2.006,74 (SPITC-1.791,74) appartenente alla β ficocianina di AFA. Gli ioni in rosso identificano la serie ionica derivante dal co-isolamento del peptide 2.014,84 (SPITC-1.799,85) di sequenza YVTYAAIAGDASVLDDR.

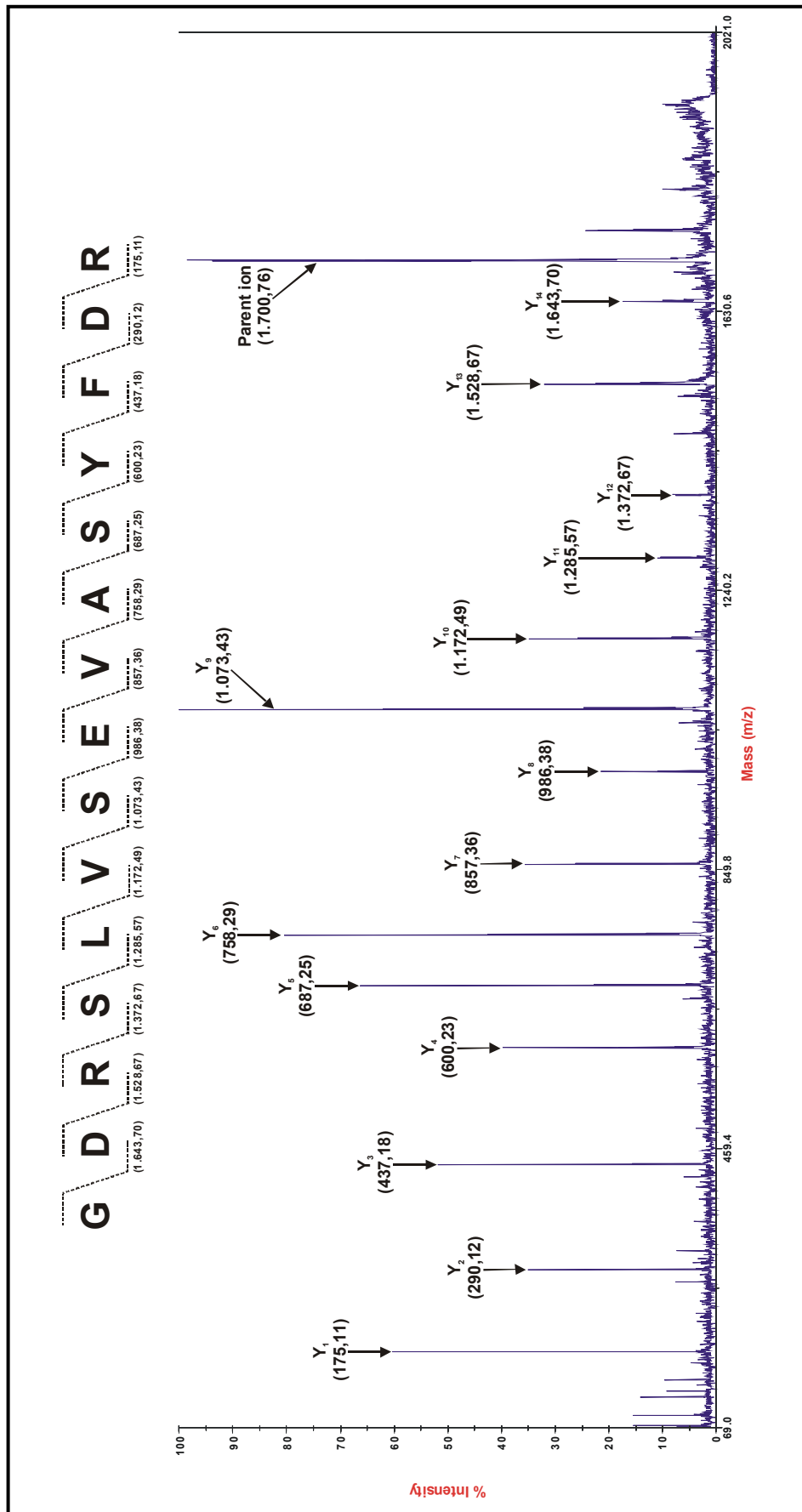


Figura 3.7 Spettro MS/MS relativo al peptide 1.915.74 (SPITC-1.700.76) appartenente alla β ficocianina di AFA.

un'altra serie ionica generata dalla frammentazione del peptide 1.799,80 Da di sequenza YVTYAAIAGDASVLDDR (2.014,80 Da derivatizzato con SPITC) che viene isolato con il primo a causa dei valori di massa molto vicini. Il secondo spettro è invece relativo al peptide 1.700,76 Da (1.915, 74 Da derivatizzato con lo SPITC) ed ha fornito la sequenza GDRSLVSEVASYFDR consentendo di elaborare due possibili sequenze per la proteina considerata.

Nostoc_sp._PCC_7120_gi 38894_	MTLDVFTKVVVSQADSRGEFLSNEQLDALANVVKEGNKRLDVVNRTSNAS
Anabaena_mendotae_gi 51105738	-----YVTYAAIAGDAS
Aphanizomenon_flos-aquae_gi 1	-----AGDAS
De_novo_sequence_1	-----GEFLSTEQLDALTAVVXX--RRLDVVNRTSNAS
De_novo_sequence_2	-----GEFLSTEQLDALTAVVXX--RRLDVVNRTSNAS
	..: **
Nostoc_sp._PCC_7120_gi 38894_	AIVTNAARALFEEQPQLIAPGGNAYTNRRMAACLRLDMEIILRYVTYAILA
Anabaena_mendotae_gi 51105738	VLDDRCNLGLRETYQALGTPGSSVAVG-----VGKMKAAIAIV
Aphanizomenon_flos-aquae_gi 1	VLDDRCNLGLRETYQALGTPGSSVAVG-----VGKMKAAIAIV
De_novo_sequence_1	AIVTNAARSLFEEQPQLIAPGGQAYTNR-----RDMEIILRYVTYAAIA
De_novo_sequence_2	AIVTNAARSLFEEQPQLIAPGGQAYTNR-----RDMEIILRYVTYAAIA
	..: .. ** * * : ** : : : .
Nostoc_sp._PCC_7120_gi 38894_	GDASVLDDRCNLGLRETYQALGTPGSSVAVGVQKMKDAAVGIANDPNGIT
Anabaena_mendotae_gi 51105738	NDPNGITK-----
Aphanizomenon_flos-aquae_gi 1	NDPNGITK-----
De_novo_sequence_1	GDASVLDD-----
De_novo_sequence_2	GDASVLDD-----
	. * . : .
Nostoc_sp._PCC_7120_gi 38894_	KGDCSQLISEVACYFDRAAAVAVG-
Anabaena_mendotae_gi 51105738	-GDCSSLVSEVASYFDRAAAAVA
Aphanizomenon_flos-aquae_gi 1	-GDCSSLVSEVASYFDRAAAVAV
De_novo_sequence_1	RGDCSSLVSEVASYFDR-----
De_novo_sequence_2	RGDRS-LVSEVASYFDR-----
	** * * : ** : * . ****

Figura 3.8 Allineamento di sequenze per l' β ficocianina.

Nostoc_sp._PCC_7120_gi 995731	MVKTPITEAIAAADTQGRFLGNTELQSARGRYERAAASLEAARGLTSNAQ
Anabaena_mendotae_gi 5110573	-MKTPITEAIASADTQGRFLSNTELQAVNGRLVRAAASMEAARGLTANAQ
Aphanizomenon_flos-aquae_gi 1	-MKTPITEAIASADTQG-----
De_novo_sequence	-MKTPITEAIASADTQGRFLSNTELQAVDG---RAAASLEAAR-----AQ
	: *****: *****
Nostoc_sp._PCC_7120_gi 995731	RLIDGATQAVYQKFPYTTTQTPGPQFAADSRGKSKCARDVGHYLRITITSL
Anabaena_mendotae_gi 5110573	KLIDSATSAVYSKFPYTTSTPRNQFAADSRGKAKCARDVGHYLR-----
Aphanizomenon_flos-aquae_gi 1	-----
De_novo_sequence	KLIDGATSAVYSKFPYTTSTPGNQYASDARDVG-----
Nostoc_sp._PCC_7120_gi 995731	VAGGTGPLDEYLIAGLAEINSTFDLSPSWYVEALKHIKANHGLSGQAANE
Anabaena_mendotae_gi 5110573	-----
Aphanizomenon_flos-aquae_gi 1	-----
De_novo_sequence	-----HYLRANHGLSGQAANE
Nostoc_sp._PCC_7120_gi 995731	ANTYIDYAINALS
Anabaena_mendotae_gi 5110573	-----
Aphanizomenon_flos-aquae_gi 1	-----
De_novo_sequence	ANTYIDYAINALS

Figura 3.9 Allineamento di sequenze per l' α ficocianina.

Entrambe queste sequenze sono state allineate per confronto con le sequenze di altre specie conosciute e con il breve tratto di catena amminoacidica di AFA presente in banca dati ed i risultati sono mostrati in Fig. 3.8. Allo stesso modo la Fig 3.9 mostra l'allineamento della sequenza ottenuta per la α ficocianina con le catene appartenenti agli stessi organismi.

CAPITOLO 4

DISCUSSIONE

Aphanizomenon flos-aquae è un'alga edibile che viene ampiamente utilizzata in molti paesi come integratore alimentare della dieta per le sue qualità terapeutiche [6, 7]. All'alto contenuto in PC sono associate le comprovate proprietà antiossidanti ed antinfiammatorie di questa microalga [8-11] ed è proprio per tale motivo che in questo studio si è voluto indagare più a fondo sulla composizione e sulla caratterizzazione di queste ficobiliproteine in AFA.

Il primo step, una volta purificate le PC a partire da un estratto secco di AFA, è stato quello di verificare il risultato del processo cromatografico di purificazione, realizzato mediante una colonna di idrossiapatite, valutando i pesi molecolari delle proteine presenti nell'eluato. L'utilizzo della spettrometria di massa MALDI-TOF ha mostrato come nessun contaminante sia rintracciabile all'interno del campione. Lo spettro in Fig. 3.1 evidenzia la presenza delle due subunità e

della ficocianina che sono state identificate attraverso il peso molecolare intatto, grazie al confronto con altre ficocianine sequenziate in diversi organismi e presenti in banca dati. In particolare i valori di massa ottenuti pari a 17.714,7 Da e 19.220,6 Da ben si accordano con i pesi molecolari delle due proteine in *Nostoc* sp. PCC 7120. Va comunque tenuto in considerazione che il pigmento ficocianobilina risulta rimanere ancorato alla proteina [25] contribuendo con 587 Da alla definizione del valore di massa. In particolare l' ficocianina lega una bilina, mentre due molecole del pigmento sono associate alla subunità , da che se ne deduce che la sequenza amminoacidica delle stesse subunità in *Aphanizomenon* risulta più corta quando confrontata con quella di *Nostoc*. La presenza di ulteriori due picchi di spalla rispetto ai segnali principali presenti nello spettro ha posto l'interrogativo sulla loro identità. Si è escluso che questi picchi possano corrispondere ad alloficocianine (APC) che sembrano essere assenti in AFA in accordo con quanto precedentemente dimostrato in un recente studio [26].

La mancanza delle APC è ulteriormente supportata dalle analisi di massa che non hanno mostrato la presenza di nessun peptide la cui sequenza potesse essere attribuita a queste ficobiliproteine o ad altri prodotti proteici diversi dalle PC. Alla luce di tutto questo, le ipotesi che sono sembrate più verosimili per spiegare la presenza di tali picchi nello spettro di massa erano due: la presenza di qualche modifica post-traduzionale a carico delle PC o la possibile esistenza di isoforme

differenti per le due subunità. Al fine di cercare una risposta a tale quesito si è proceduto ad una caratterizzazione delle due proteine mediante un'indagine proteomica approfondita utilizzando un approccio di tipo "bottom-up".

La Fig. 3.2 mostra il risultato della separazione elettroforetica delle PC purificate che migrano in due bande distinte in prossimità dei 20 kDa. Le bande sono state escisse, sottoposte a digestione proteolitica in-gel con tripsina e le miscele di peptidi ottenute sono state analizzate mediante spettrometria di massa tandem utilizzando un MALDI-TOF/TOF. Le ricerche in banca dati utilizzando sia i dati relativi al "peptide mass fingerprinting" sia i dati di sequenza hanno permesso di identificare le due subunità α e β della ficocianina seppure, come è evidente da quanto si vede in Tabella 3.3, solo pochi peptidi sono stati sequenziati, risultando in un basso valore di copertura di sequenza. Nell'intento, quindi, di ottimizzare i risultati di questa ricerca si è proceduto all'applicazione di un metodo di derivatizzazione chimica che fa uso del reagente SPITC, per ottenere un numero maggiore di spettri di buona qualità da poter sequenziare de novo. L'utilizzo dello SPITC, oltre a garantire l'ottenimento di una migliore frammentazione che aumenta il segnale delle serie ioniche, in particolare della serie y che facilita l'interpretazione degli spettri (Fig. 3.3 e Fig. 3.5), ha dimostrato anche di agire incrementando l'efficienza di ionizzazione dei peptidi in sorgente. Alcuni ioni, indicati dall'asterisco in Tabella 3.5, sono presenti nello spettro di massa dopo derivatizzazione ma non compaiono invece nello spettro di massa della miscela peptidica non derivatizzata. La frammentazione di questi peptidi ha però fornito importanti dati di sequenza che sarebbero andati altrimenti persi senza l'ausilio dello SPITC. I risultati delle analisi sui peptidi indagati sono riassunti in Tabella 3.4 e 3.5. Si può notare come alcune volte dei peptidi sequenziati, indicati con ‡, non siano propriamente dei peptidi triptici completi ma derivino da un'ulteriore frammentazione di ioni di dimensioni maggiori. Tale frammentazione è forse indotta dalla derivatizzazione stessa considerando che tali frammenti non sono presenti negli spettri delle proteine non derivatizzate seppure non legano lo SPITC.

Le Fig. 3.6 e 3.7 mostrano invece un risultato molto interessante evidenziando due diverse sequenze corrispondenti ad un tratto della catena della PC prossimo all'estremità C-terminale della proteina. Le due sequenze, GDCSSLVSEVASYFDR relativa al peptide di 1.791,74 Da (2.006,74 Da derivatizzato con lo SPITC) e GDRSLVSEVASYFDR relativa al peptide di 1.700,76 Da (1.915,74 Da derivatizzato con lo SPITC), avvalorano l'ipotesi dell'esistenza di

diverse isoforme per le due catene della ficocianina. La prima di queste sequenze è comune a molti altri microorganismi, come si può vedere dagli allineamenti di sequenza mostrati in Fig. 3.8, mentre la seconda sembra essere peculiare di *Aphanizomenon*. Purtroppo la differenza di 91 Da tra le due sequenze non è sufficiente da sola a spiegare la differenza di circa 165 Da tra il peso molecolare della proteina intatta, rilevato nello spettro di massa, e quella del suo picco di spalla. Considerando che la copertura di sequenza per il polipeptide è di circa 60%, si può ipotizzare che ulteriori differenze possano essere presenti in altre porzioni della catena polipeptidica che sono sfuggite all'analisi di massa. In maniera analoga, si ipotizza che esistano delle differenze che non siano state rilevate anche a carico della catena dell' PC che giustifichino l'esistenza di isoforme differenti. In quest'ottica, ciò che ci si propone di fare nel prosieguo di questo studio è di utilizzare enzimi proteolitici differenti per una caratterizzazione più estesa delle PC e di applicare tecniche come l'IEF per confermare l'esistenza di isoforme differenti. Le possibili isoforme, infatti, non possono essere risolte elettroforeticamente utilizzando una separazione in base al peso molecolare a causa della piccola differenza nei valori di massa ma potrebbero mostrare significative alterazioni del pI tali da permettere di distinguerle in maniera netta.

CAPITOLO 5

BIBLIOGRAFIA

- [1] Bryant, D. A. (1986) The Cyanobacterial Photosynthetic Apparatus: Comparison to those of Higher Plants and Photosynthetic Bacteria. In *"Photosynthetic Pikoplankton"* T. Platt, W. K. W. Li, (Eds.) pp. 423-500. Canadian Bulletin of Fisheries and Aquatic Sciences, Vol. 214, Department of Fisheries and Oceans, Ottawa, Ontario, Canada.
- [2] Jensen, G. S., Ginsberg, D. I., Huerta, P., Citton, M., Drapeau, C. (2000) Consumption of *Aphanizomenon flos-aquae* has rapid effects on the circulation and function of immune cells in humans. *Journal of the American Nutraceutical Association* **2**: 50–58.
- [3] Jensen, G. S., Ginsberg, D. I., Drapeau, C. (2001) Blue-green algae as an immuno-enhancer and biomodulator. *Journal of the American Nutraceutical Association* **3**: 24–30.
- [4] Saker, M. L., Jungblut, A.-D., Neilan, B. A., Rawn, D. F. K., Vasconcelos, V. M. (2005) Detection of microcystin synthetase genes in health food supplements containing the freshwater cyanobacterium *Aphanizomenon flos-aquae*. *Toxicon* **46**: 555-562.
- [5] Chen, Y., Liu, J., Yang, W. (2003) Effect of *Aphanizomenon flos-aquae* toxins on some blood physiological parameters in mice *Wei Sheng Yan Jiu* **32**: 195-197.
- [6] Pugh, N., Pasco, D. S. (2001) Characterization of human monocyte activation by a water soluble preparation of *Aphanizomenon flos-aquae*. *Phytomedicine* **8**: 445–453.
- [7] Pugh, N., Ross, S. A., ElSohly, H. N., ElSohly, M. A., Pasco, D. S. (2001) Isolation of three high molecular weight polysaccharide preparations with potent immunostimulatory activity from *Spirulina platensis*, *Aphanizomenon flos-aquae* and *Chlorella pyrenoidosa*. *Planta Med.* **67**: 737–742.

- [8] Bhat, V. B., Madyastha, K. M. (2000) C-phycoerythrin: a potent peroxyl radical scavenger *in vivo* and *in vitro*. *Biochem. Biophys. Res. Commun.* **275** (1): 20–25.
- [9] Romay, C., Gonzalez, R. (2000) Phycocyanin is an antioxidant protector of human erythrocytes against lysis by peroxyl radicals. *J. Pharm. Pharmacol.* **52** (4): 367–368.
- [10] Romay, C., Armesto, J., Ramirez, D., Gonzalez, R., Ledon, N., Garcia, I. (1998) Antioxidant and anti-inflammatory properties of C-phycoerythrin from blue-green algae. *Inflamm. Res.* **47** (1): 36–41.
- [11] Reddy, C. M., Bhat, V. B., Kiranmai, G., Reddy, M. N., Reddanna, P., Madyastha, K. M. (2000) Selective inhibition of cyclooxygenase-2 by C-phycoerythrin, a biliprotein from *Spirulina platensis*. *Biochem. Biophys. Res. Commun.* **277** (3): 599–603.
- [12] Stocker, R., McDonagh, A. F., Glazer, A. N., Ames, B. N. (1990) Antioxidant activities of bile pigments: biliverdin and bilirubin. *Methods Enzymol.* **186**: 301–309.
- [13] Bhat, V. B., Madyastha, K. M. (2001) Scavenging of peroxynitrite by phycocyanin and phycocyanobilin from *Spirulina platensis*: protection against oxidative damage to DNA. *Biochem. Biophys. Res. Commun.* **285** (2): 262–266.
- [14] Benedetti, S., Benvenuti, F., Pagliarani, S., Francogli, S., Scoglio, S., Canestrari, F. (2004) Antioxidant properties of a novel phycocyanin extract from the blue-green alga *Aphanizomenon flos-aquae* *Life Sci.* **75**: 2353–2362.
- [15] Yoshida, A., Takagaki, Y., Nishimune, T. (1996) Enzyme immunoassay for phycocyanin as the main component of spirulina color in foods. *Biosci. Biotechnol. Biochem.* **60** (1): 57–60.
- [16] Cohen, Z. (1986) Products from microalgae. In: “*Handbook of Microalgal Mass Culture*” Richmond, A. (Ed.),. CRC Press, Boca Raton, pp. 421–454.

- [17] Glazer, A. N., Stryer, L. (1990) Phycobiliprotein-avidin and phycobiliprotein-biotin conjugates. *Methods Enzymol.* **184**: 188–194.
- [18] MacColl, R. (1998) Cyanobacterial phycobilisomes. *J. Struct. Biol.* **124**: 311–334.
- [19] MacColl, R. (1983) Stability of allophycocyanin's quaternary structure. *Arch. Biochem. Biophys.* **223**: 24–32.
- [20] MacColl, R., Berns, D., Koven, N.L. (1971) Effect of salts on C-phycocyanin. *Arch. Biochem. Biophys.* **146**: 477–482.
- [21] MacColl, R., Guard Mac-Friar, D. (1987) “*Phycobiliproteins*”, pp.141–169. CRC press Inc., Boca Raton, FL, USA.
- [22] Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4 *Nature.* **227**: 680–685.
- [23] Shevchenko A., Wilm M., Vorm O., Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* **68**: 850–858.
- [24] Gobom J., Nordhoff E., Mirgorodskaya E., Ekman R., Roepstorff P. (1999) Sample purification and preparation technique based on nano-scale reversed-phase columns for the sensitive analysis of complex peptide mixtures by matrix-assisted laser desorption/ionization mass spectrometry. *J. Mass Spectrom.* **34**: 105–116.
- [25] L. Zolla, M. Bianchetti, (2001) High-performance liquid chromatography coupled on-line with electrospray ionization mass spectrometry for the simultaneous separation and identification of the *Synechocystis* PCC 6803 phycobilisome proteins *J. Chromatogr. A* **912**: 269–279.

- [26] Benedetti, S., Rinalducci, S., Benvenuti, F., Francogli, S., Pagliarani, S., Giorgi, L., Micheloni, M., D'Amici, G. M., Zolla, L., Canestrari, F. (2005) Purification and characterization of phycocyanin from the blue-green alga *Aphanizomenon flos-aquae*. *J. Chromatogr. B*

RINGRAZIAMENTI

Al termine di questo lavoro di tesi desidero ringraziare tutti coloro che sono stati presenti in questi anni e che mi hanno permesso di compiere quest'esperienza.

Ringrazio la mia famiglia: mia madre Anna e mio padre Mario per il loro supporto e per il loro amore incondizionato e continuo, cosa che suona in maniera retorica finchè uno "davvero" non si ferma a pensare a quanto sia grande. Voglio dire grazie a mio fratello Roberto, a sua moglie Teodora e a mio fratello Carlo.

Ringrazio il Prof. Lello Zolla per la sua supervisione, i suoi consigli e per la sua vulcanica fervidità di idee che mi ha stimolato nel compiere questo cammino. Un grazie sentito va alla Dr. ssa Anna Maria Timperio per la sua costante presenza ed assistenza. Grazie alla Dr. ssa Sara Rinalducci che è stata in questi anni non solo una preziosa collega ma soprattutto un grande amico con la quale ho condiviso molto imparando tanto. Mi aspetto "grandi" cose da te! Ringrazio poi tutti coloro con i quali ho lavorato e che mi hanno arricchito in questi anni di ricerca svolti nel laboratorio di "Proteomica e Genomica": la Dr. ssa Maria Bianchetti, la Dr. ssa Luisiana Lattanzi, la Dr. ssa Sonia Troiani, la Dr. ssa Deborah Proietti Petretti, il Dr. Paolo Catarcini, il Dr. GianMaria D'Amici.

Special thanks to my supervisors abroad Prof. Eva-Mari Aro and Prof. Peter Roepstorff who gave me the great chance to visit their labs and deepen my knowledges in science. Thanks to Dr. Shabaz Mohammed, Dr. Miren Josu Omaetxebarria and Dr. Emmanuelle Sachon for their true friendship and precious advice. Thanks to all the members of PR-group who had a high level of patience in teaching me what MS is for!!

Grazie ai miei amici Maria, Francesco, Umberto, Salvatore, Annarita, Serena, Flavia e Simona.

Voglio anche ringraziare tutti gli afferenti al Dipartimento di Protezione delle Piante: il Prof. Paolo Magro, il Prof. Gabriele Chilosi, il Dr. Roberto Reda, la Dr. ssa Maria Pia Aleandri e Diana Martignoni.

Grazie a tutti coloro che non ho potuto citare ma che sono costantemente presenti nel mio cuore e nei miei pensieri. Grazie a tutti coloro che hanno reso ciò possibile, che mi hanno ispirato, che mi hanno fatto crescere, che mi hanno arricchito, che mi hanno aiutato, che mi hanno portato fin qui a dire GRAZIE!

PUBBLICAZIONI

- Ciambella, C., Roepstorff, P., Aro, E.-M., Zolla, L. (2005) A proteomic approach for investigation of photosynthetic apparatus in plants. *Proteomics*, **5**: 746-757.
- Gil-Agusti, M. T., Campostrini, N., Zolla, L., Ciambella, C., Invernizzi, C., Righetti, P. G. (2005) Two-dimensional mapping as a tool for classification of green coffee bean species. *Proteomics*, **5**: 710-718.

COMUNICAZIONI A CONGRESSI

- Zolla, L., Timperio, A. M., Rinalducci, S., Ciambella, C., Catarcini, P., Huber, C.G. Separation and identification of photosynthetic membrane proteins on accurate intact molecular mass determination. *"IV Corso di Spettrometria di Massa nello Studio Integrato del Genoma e Proteoma" Vitorchiano (VT) 9-12 Dicembre 2002.*
- Huber, C. G., Timperio, A. M., Rinalducci, S., Catarcini, P., Ciambella, C., Zolla, L. Liquid Chromatography-Electrospray Ionization Mass Spectrometry in Proteomic and Genomic Research. *"XXI Congresso Nazionale della Società Chimica Italiana" Torino 22-27 Giugno 2003.*
- Ciambella, C., Rinalducci, S., Timperio, A. M., D'Amici, G. M., Aro, E.-M., Roepstorff, P., Zolla, L. Proteomics of thylakoid membranes from barley by 2D-BN/SDSPAGE and tandem ESI-MS/MS. *"Proteine 2004" Viterbo 20-22 Maggio 2004.*
- Rinalducci, S., Timperio, A. M., Ciambella, C., Huber, C. G., Righetti, P. G., Zolla, L. Intact mass measurements for unequivocal identification of hydrophobic photosynthetic PSI and PSII antenna proteins. *"Italian Proteome Society, first congress" Verona 27-29 Maggio 2004.*

- Timperio, A. M., Rinalducci, S., Ciambella, C., D'Amici, G. M., Huber, C. G., Zolla, L. Proteomics of photosystem I protein components in different plant species performed by HPLC on line with ESI-MS. *"Italian Proteome Society, first congress" Verona 27-29 Maggio 2004.*

- Timperio, A. M., Huber, C. G., Rinalducci, S., Ciambella, C., Zolla, L. Separation and identification of the light harvesting proteins contained in grana and stroma thylakoid membrane fractions. *"Italian Proteome Society, first congress" Verona 27-29 Maggio 2004.*

- Gil-Agusti, M. T., Campostrini, N., Zolla, L., Ciambella, C., Invernizzi, C., Righetti, P.G. Classification of green coffee beans by differential proteomic analysis via two-dimensional maps *"Italian Proteome Society, first congress" Verona 27-29 Maggio 2004.*

- Romero-Puertas, M.C., Campostrini, N., Perazzolli, M., Zolla, L., Ciambella, C., Righetti, P.G., Delledonne, M. Proteomic study of plant-specific proteins S-nitrosylated in response to bacterial infection. *"Italian Proteome Society, first congress" Verona 27-29 Maggio 2004.*

- Ciambella, C., Roepstorff, P., Rinalducci, S., Timperio, A. M. D'Amici, G. M., Aro, E.-M., Zolla, W. Proteomics of thylakoid membrane proteins from iron-deficient spinach plants. *"Second IPSO Congress – Proteomics and Genomics" Viterbo 29 Maggio – 1 Giugno 2005.*

- Zaninotto, F. Romero-Puertas, M. C., Campostrini, N., Perazzolli, M., Matté, A., Degliantoni, S., Zolla, L., Ciambella, C., Righetti, P. G., Delledonne, M. Analysis of S-nitrosylated proteins in *Arabidopsis thaliana*. *"Second IPSO Congress – Proteomics and Genomics" Viterbo 29 Maggio – 1 Giugno 2005.*

- Rinalducci, S., Timperio, A. M., Ciambella, C., D'Amici, G. M., Zolla, L. Proteomic analysis of photosystem I components from different plant species. *"Second IPSO Congress – Proteomics and Genomics" Viterbo 29 Maggio – 1 Giugno 2005.*

• Ciambella, C., Romero-Puertas, M. C., Campostrini, N., Perazzolli, M., Righetti, P.G., Rinalducci, S., D'Amici, G. M., Fagioni, M., Timperio, A. M., Delledonne, M., Roepstorff, P., Zolla, L. Proteome of S-nitrosylated proteins expressed in *Arabidopsis thaliana* in response to pathogen infection. *"MASSA 2005 An International Symposium on Mass Spectrometry "Roma Università Cattolica, 28 Giugno - 1 Luglio 2005.*

• D'Amici, G. M., Rinalducci, S., Antonioli, P., Ciambella, C., Fagioni, M., Timperio, A. M., Righetti, P. G., Zolla, L. Protein pre-fractionation by three-phase system of thylakoid membrane proteins for facilitated 2D-IEF/SDS-PAGE and MS identification. *"MASSA 2005 An International Symposium on Mass Spectrometry "Roma Università Cattolica, 28 Giugno - 1 Luglio 2005.*