UNIVERSITA' DEGLI STUDI DI VERONA
FACOLTA' DI SCIENZE MM.FF.NN.
DIPARTIMENTO DI BIOTECNOLOGIE

DOTTORATO DI RICERCA IN
BIOTECNOLOGIE MOLECOLARI, INDUSTRIALI E AMBIENTALI
XXI CICLO

2D-PAGE COUPLED TO MASS SPECTROMETRY FOR PROTEOMIC ANALYSIS OF HUMAN, MICROBIAL AND PLANT SAMPLES

S.S.D. BIO/10

Coordinatore: Prof. HUGO LUIS MONACO

Tutor: Prof. PIER GIORGIO RIGHETTI

Dott.ssa DANIELA CECCONI

Dottorando: Dott. ALBERTO MILLI

24 Aprile 2009

A te

“To truth itself I gave what is false no less than what is true as its rightful province, and showed that the false and the true are merely forms of intellectual existence.”

Oscar Wilde - *De Profundis*

“Alla verità stessa concessi, come le spettava di diritto, il falso non meno del vero e mostrai come falso e vero sono solamente mere forme del nostro intelletto.”

Oscar Wilde - *De Profundis*

UNIVERSITÁ DEGLI STUDI DI VERONA
DIPARTIMENTO DI BIOTECNOLOGIE
Strada Le Grazie n. 15, Ca' Vignal – 37134 Verona (Italia)
Tel. + 39 045 802 7945 (ordini), 7934 (fatture) – Fax + 39 045 8027929 – E-mail: segreteria-dst@sci.univr.it
Codice Fiscale 93009870234 – Partita IVA 01541040232

Alla cortese attenzione del Collegio dei Docenti:

Oggetto: profilo del Dott. Alberto Milli, dottorando di ricerca in Biotecnologie Molecolari, Industriali e Ambientali, XXI ciclo, Dipartimento di Biotecnologie.

Il Dott. Milli, nel corso del triennio di attività scientifica svolta nell'ambito del Dottorato di Ricerca in Biotecnologie Molecolari, Industriali e Ambientali presso il Dipartimento di Biotecnologie dell'Università di Verona, ha condotto un brillante ed approfondito lavoro inerente gli aspetti connessi all'analisi proteomica comparativa mediante elettroforesi bidimensionale accoppiata a spettrometria di massa di campioni umani, microbici e vegetali.
Nello svolgimento del lavoro di ricerca il Dott. Milli ha affrontato le problematiche connesse allo studio del proteoma del tumore al colon-retto per valutare gli effetti molecolari del trattamento con un nuovo tipo di inibitore delle istone deacetilasi, utilizzando in particolare due diversi metodi statistici per l'analisi dei profili proteici.
Sono state affrontate inoltre le tematiche relative la preparazione e la separazione dei campioni proteici, allo scopo di valutare gli effetti dell'acido tannico su un ceppo del microorganismo del vino *Lactobacillus plantarum* e le variazioni nell'espressione proteica alla base della suscettibilità della vite all'agente della Peronospora, entrambi aspetti molto importanti nella produzione vitivinicola.
I temi sono stati affrontati con rigore scientifico, conducendo egregie analisi del proteoma di campioni umani, batterici e vegetali e portando alla dettagliata spiegazione delle dinamiche di proteine coinvolte negli eventi riguardanti le suddette problematiche di interesse biomedico ed economico-ambientale.
La produzione scientifica svolta dal Dott. Milli in questi tre anni, in seguito ai progetti attuati, ha portato alla pubblicazione dei seguenti articoli:

1. Proteomic analysis of Oenococcus oeni freeze-dried culture to assess the importance of cell acclimation to conduct malolactic fermentation in wine. Cecconi D, Milli A, Rinalducci S, Zolla L, Zapparoli G. *Electrophoresis*, 2009, submitted.
2. Effect of tannic acid on Lactobacillus plantarum wine strain during starvation: A proteomic study. Cecconi D, Cristofoletti M, Milli A, Antonioli P, Rinalducci S, Zolla L, Zapparoli G. *Electrophoresis*, 2009,30, 957-965.
3. Signal transduction pathways of mantle cell lymphoma: A phosphoproteome-based study. Cecconi D, Zamò A, Bianchi E, Parisi A, Barbi S, Milli A, Rinalducci S, Rosenwald A, Hartmann E, Zolla L, Chilosi M. *Proteomics*, 2008, 21, 4495-4506.
4. A proteomic approach for evaluating the cell response to a novel histone deacetylase inhibitor in colon cancer cells. Milli A, Cecconi D, Campostrini N, Timperio AM, Zolla L,

UNIVERSITÁ DEGLI STUDI DI VERONA
DIPARTIMENTO DI BIOTECNOLOGIE
Strada Le Grazie n. 15, Ca' Vignal – 37134 Verona (Italia)
Tel. + 39 045 802 7945 (ordini), 7934 (fatture) – Fax + 39 045 8027929 – E-mail: segreteria-dst@sci.univr.it
Codice Fiscale 93009870234 – Partita IVA 01541040232

Righetti SC, Zunino F, Perego P, Benedetti V, Gatti L, Odreman F, Vindigni A, Righetti PG. *Biochim. Biophys. Acta*, 2008, 11, 1702-1710.

5. Application of Partial Least Squares Discriminant Analysis and variable selection procedures: a 2D-PAGE Proteomic-Based Study. Marengo E, Robotti E, Bobba M, Milli A, Campostrini N, Righetti SC, Cecconi D, Righetti PG. *Anal. Bioanal. Chem.,* 2008, 5, 1327-1342.
6. Proteomic analysis of rat hippocampus and frontal cortex after chronic treatment with fluoxetine or putative novel antidepressants: CRF1 and NK1 receptor antagonists. Carboni L, Vighini M, Piubelli C, Castelletti L, Milli A, Domenici E. *Eur. Neuropsychopharmacology,* 2006, 16, 521-537.
7. Searching for markers of Creutzfeldt-Jakob disease in cerebrospinal fluid by two-dimensional mapping. Piubelli C, Fiorini M, Zanusso G, Milli A, Fasoli E, Monaco S, Righetti PG. *Proteomics,* 2006, 6, S256-261.

Il numero delle pubblicazioni prodotte certifica l'importante impegno profuso dal Dott. Milli nell'adempimento del suo lavoro.
L'elaborato della tesi è sicuramente collocabile al livello di ottimo.

I responsabili Scientifici
Prof. Pier Giorgio Righetti
Dott.ssa Daniela Cecconi

Visto
Il Coordinatore del Dottorato di Ricerca
in Biotecnologie Molecolari, Industriali e Ambientali
Prof. Hugo Monaco

Verona, 24 aprile 2009

RIASSUNTO

Il progetto di dottorato è stato condotto nel Laboratorio di Proteomica del Dipartimento di Biotecnologie dell'Università di Verona. Per la stesura di tale progetto sono state instaurate collaborazioni con alcuni laboratori, sia interni allo stesso Dipartimento sia appartenenti ad altre Università o Istituti di Ricerca. In particolare, per quanto riguarda lo studio su cellule di cancro colorettale trattate con un nuovo chemiofarmaco, il Laboratorio del Prof. Zunino (Fondazione IRCCS, Istituto Nazionale dei Tumori, Milano) ha fornito tutti i campioni biologici, mentre il gruppo del Prof. Marengo (Dipartimento di Scienze dell'Ambiente e della Vita, Università degli Studi del Piemonte Orientale, Alessandria) ha effettuato l'analisi statistica multivariata dei dati.

Per quanto riguarda invece la caratterizzazione microbiologica, biochimica e proteomica del ceppo tannasi-positivo VP08 di *Lactobacillus plantarum*, il lavoro è stato condotto in collaborazione con il Dottor Zapparoli (Laboratorio di Microbiologia, Dipartimento di Biotecnologie, Università di Verona).

L'indagine proteomica su campioni di foglie di *Vitis vinifera* suscettibile a *Plasmopara viticola* è stata realizzata in collaborazione col gruppo della Dottoressa Polverari (Laboratorio di Biotecnologie Fitopatologiche, Dipartimento di Biotecnologie, Università di Verona).

Infine, il gruppo del Prof. Zolla (Laboratorio di Proteomica, Dipartimento di Scienze Ambientali, Università degli Studi della Tuscia, Viterbo) ha effettuato le analisi di spettrometria di massa per l'identificazione delle proteine di interesse individuate nei diversi studi. Ulteriori analisi di massa sono state svolte, in parte, presso il medesimo Laboratorio di Proteomica del Dipartimento di Biotecnologie dell'Università di Verona e, in parte, presso il Laboratorio di Proteomica del Dottor Vindigni (Centro Internazionale di Ingegneria Genetica e Biotecnologia "ICGEB", Trieste).

In questo progetto di dottorato sono stati applicati i metodi classici della proteomica comparativa basata sull'elettroforesi bidimensionale per l'analisi di campioni umani, microbici e vegetali legati a tre differenti problematiche, riguardanti rispettivamente:

1) la risposta di una linea cellulare di cancro al colon-retto ad un nuovo tipo di inibitore delle istone deacetilasi

2) l'effetto dell'acido tannico sul microrganismo del vino Lactobacillus plantarum *in condizioni di carenza di nutrienti*

3) i cambiamenti nel proteoma di foglia di Vitis vinifera *cultivar Pinot Noir a diversi tempi dopo infezione con* Plasmopara viticola

Per quanto riguarda la prima parte del progetto, l'attenzione è stata focalizzata su una linea cellulare tumorale di cancro colorettale (CRC). Il CRC è uno dei più comuni tumori maligni, che fa contare almeno un milione di nuovi casi in tutto il mondo e porta a più di 500000 morti ogni anno. Circa la metà dei pazienti con cancro del colon-retto sviluppa metastasi durante il decorso della malattia. Come tutti i tumori, il CRC risulta essere il risultato di eventi genetici ed epigenetici. Per quanto riguarda le alterazioni epigenetiche, è noto che l'ipoacetilazione degli istoni causa una repressione trascrizionale. L'acetilazione/deacetilazione degli istoni è controllata dall'azione degli enzimi istone-acetil-trasferasi (HAT) e istone-deacetilasi (HDAC). Per questa ragione, è stato proposta la modulazione della repressione genica a livello epigenetico come un approccio interessante per il controllo della crescita della massa tumorale e le HDAC sono riconosciuti come potenziali target per questo approccio. Lo scopo di questa parte di progetto di dottorato è stato investigare l'effetto molecolare di un inibitore delle HDAC ad ampio spettro (l'RC307) sul profilo proteico della linea cellulare HCT116 di CRC. Le cellule della linea HCT116 sono state messe in coltura con e senza RC307 e sono stati settati i protocolli per estrarre le proteine sia dei lisati totali, sia degli estratti nucleari. Mediante elettroforesi bidimensionale dei campioni proteici e l'analisi delle immagini tramite il software PDQuest è stato possibile individuare, dopo il trattamento con il farmaco, un totale di 48 e 46 proteine differenzialmente espresse, nei lisati totali e negli estratti nucleari rispettivamente. L'identificazione mediante spettrometria di massa nanoHPLC-ESI-IT e MALDI-TOF-TOF ha rivelato che tra le proteine modulate molte sono correlate a vari processi cellulari che coinvolgono ad esempio la proliferazione, il ciclo cellulare e la regolazione apoptotica, l'espressione genica e l'organizzazione della cromatina e del citoscheletro e alcune delle variazioni più interessanti sono state validate mediante western blot. Nel complesso questo lavoro ha permesso di

dimostrare come l'inibitore delle HDAC RC307 sia efficace contro la crescita e lo sviluppo del cancro al colon-retto. Il farmaco regola la repressione genica in modo epigenetico, portando all'inibizione della proliferazione cellulare e alla morte per apoptosi mediante la regolazione di proteine chiave nella regolazione di questi processi. Per verificare la veridicità dei risultati ottenuti dal confronto dei profili proteici effettuato con il PDQuest, si è infine proceduto con l'analisi multivariata degli stessi set di dati di lisati e nuclei mediante un tipo di analisi statistica multivariata, il Partial Least Squares – Discriminant Analysis (BE PLS-DA). Questa ulteriore analisi ha permesso di validare i dati ottenuti tramite lo studio al PDQuest, ricavando informazioni aggiuntive. Infatti il confronto tra i due tipi di analisi ha messo in luce che l'analisi differenziale mediante il t-test (utilizzato dal PDQuest) identifica un sotto-insieme degli spot significativi identificati dalla BE PLS-DA, che risulta essere un metodo statistico più robusto in grado di individuare variazioni tra loro correlate.

La seconda parte della tesi riguarda la caratterizzazione microbiologica, biochimica e proteomica del ceppo tannasi positivo VP08 di *Lactobacillus plantarum.* I tannini vegetali sono presenti nel suolo, nelle derrate alimentari e, come conseguenza, nei tratti digerenti degli uomini e degli animali. L'acido tannico (TA), un tannino idrolizzabile, ha un effetto inibitorio sulla crescita di alcuni batteri, compresi i batteri lattici (LAB). Alcuni batteri sono in grado di degradare i tannini grazie alla produzione dell'enzima tannin acyl idrolasi, comunemente chiamata tannasi. Tra le specie di LAB di interesse enologico, l'unica che risulta essere tannasi-positiva è *L. plantarum*. In questa parte del progetto di dottorato, è stato indagato l'effetto fisiologico e molecolare dell'acido tannico sul ceppo del vino tannasi-positvo di *L. plantarum* VP08, mediante saggi enzimatici e di crescita e mediante l'analisi proteomica differenziale del batterio cresciuto in presenza e assenza di TA. Le cellule cresciute in presenza di TA hanno mostrato una maggiore attività della enzima tannasi rispetto a quelle cresciute in glucosio. Le analisi cinetiche di crescita hanno inoltre dimostrato che in presenza di TA la popolazione cellulare rimane in fase stazionaria più a lungo rispetto alle cellule cresciute in glucosio e che, anche in presenza di un'alta concentrazione di tannini, la popolazione cellulare riesce ad adattarsi (possibilmente per effetto dell'induzione della tannasi), portando successivamente ad un incremento della probabilità di sopravvivenza. É stata quindi

effettuata un'analisi proteomica comparativa mediante 2D-PAGE per individuare come si esplicita a livello molecolare l'aumento di sopravvivenza indotto dal TA. L'analisi al PDQuest delle immagini delle mappe proteiche ha permesso di individuare un totale di 36 spot differenzialmente espresse (di cui 15 sotto-espresse e 21 sovra-espresse) nel campione di cellule cresciute in presenza di TA rispetto alle cellule cresciute in glucosio. Le proteine identificate mediante analisi di spettrometria di massa sono risultate essere coinvolte nella glicolisi, nel metabolismo degli aminoacidi, nella traduzione e nel ripiegamento delle proteine. I risultati suggeriscono come la presenza di TA non porti ad un cambiamento globale a livello di espressione proteica nelle cellule batteriche, ma piuttosto alteri il livello di un numero relativamente basso di proteine appartenenti ad importanti processi cellulari e metabolici. Ad esempio, interessante da notare è il fatto che tutti gli enzimi coinvolti nel processo glicolitico risultano aumentati. Questo può indicare che il ceppo tannasi-positivo VP08 di *L. plantarum*, in carenza di nutrienti, è in grado di aumentare la sua sopravvivenza ottenendo energia supplementare mediante l'idrolisi tannasi-dipendente dell'acido tannico nei monomeri costitutivi (acido gallico e glucosio) e mediante la successiva sovra-espressione di enzimi glicolitici. Inoltre i dati mostrano che nelle cellule cresciute in TA vi è una sovra-regolazione di proteine coinvolte nella sintesi proteica e di alcune delle proteine che partecipano al corretto ripiegamento delle proteine e alla degradazione di quelle instabili. È possibile quindi inferire che le cellule cresciute in presenza di TA siano caratterizzate da un'aumentata attività del macchinario di ripiegamento e turn-over delle proteine, in grado di permettere l'adattamento alle condizioni di stress da carenza di nutrienti e una maggiore capacità di sopravvivenza. In conclusione questo studio ha permesso per la prima volta di indagare i meccanismi molecolari coinvolti nel rallentamento della crescita cellulare e nell'allungamento della sopravvivenza indotti dal TA in cellule di *L. plantarum* in condizioni di carenza di nutrienti, quali sono quelle che possono presentarsi naturalmente nel vino.

L'ultima parte del progetto di dottorato ha riguardato lo studio del proteoma di foglia di *Vitis vinifera* cultivar Pinot Noir a diversi tempi dopo infezione con *Plasmopara viticola.* La peronospora è una delle più gravi patologie della vite coltivata e ancora oggi rappresenta la malattia più distruttiva in Europa e negli Stati Uniti orientali. *P. viticola*,

l'agente causale della malattia, colpisce tutti gli organi verdi della pianta che presentano gli stomi e in condizioni climatiche favorevoli ed in assenza di misure di controllo può causare una distruzione del 50-75% del raccolto stagionale. D'altra parte la peronospora comporta un alto impatto ambientale, a causa del continuo impiego di fungicidi che continua tuttora. Per investigare l'effetto molecolare dell'infezione di *Plasmopara viticola* sul profilo proteico della cultivar suscettibile Pinot Noir sono quindi state messe a punto le condizioni di estrazione e di preparazione delle proteine da tessuti fogliari e sono state ottenute le mappe 2D dei campioni dopo 24, 48 e 96 ore dall'inoculazione col patogeno e dei rispettivi controlli trattati con acqua. L'analisi al PDQuest ha permesso di rivelare 53, 14 e 47 proteine differenzialmente espresse, rispettivamente a 24h, 48h e 96h dopo l'infezione. Tramite l'analisi di spettrometria di massa, sono stati identificati in totale 72 diversi polipeptidi. Le proteine regolate dall'infezione del patogeno che sono state identificate, risultano essere correlate a vari processi cellulari che coinvolgono ad esempio la fotosintesi e il metabolismo primario, la risposta a stimoli e di difesa, la regolazione del potenziale redox e del ripiegamento delle proteine. In particolare, è stato dimostrato che il patogeno porta ad una generale sotto-espressione delle proteine coinvolte nei processi di fotosintesi e ad un concomitante aumento nell'espressione degli enzimi correlati con il metabolismo dei carboidrati, in accordo con precedenti studi di trascrittomica condotti nel laboratorio di Biotecnologie Fitopatologiche della Dott.ssa Polverari. Nel complesso, questi dati, assieme a quelli che potranno essere raccolti dall'analisi su campioni di una cultivar resistente, possono permettere di delineare una strategia per l'ingegnerizzazione di piante di vite, al fine di renderle in grado di difendersi efficacemente dalla malattia.

Concludendo, tutti gli studi di analisi proteomica comparativa, effettuati durante questo progetto di dottorato, confermano che l'elettroforesi bidimensionale, accoppiata all'analisi tramite spettrometria di massa, è una tecnica efficace ed universale per l'analisi globale della variazione dell'espressione proteica e per lo studio e la comprensione dei meccanismi molecolari che stanno alla base di importanti processi di interesse biomedico, economico ed ambientale.

PREFACE

The proteome of a cell or an organelle provides information about the ensemble of proteins and protein isoforms expressed in that cell or organelle under specific physiological conditions and at a specific time. Proteomic approaches provide several novel possibilities to address biological questions. In fact, the large-scale screening approach of proteomics enables protein expression studies that are impossible to perform using classical molecular biology and biochemical techniques, in which the expression of only one or a few proteins is studied at a time. Instead, proteomic techniques allow for the analysis of up to thousands of proteins simultaneously, in any tissue or organelle, under any given physiological condition. Thus, proteomic applications are growing in many areas of research and proteomic approaches are nowadays widely exploited for cancer, microbial, and plant investigations.

The present work was focused on three proteomics studies:

1) evaluation of the cell response to a novel histone deacetylase inhibitor in colon cancer cell

2) effect of tannic acid on lactobacillus plantarum *wine strain during starvation*

3) analysis of grapevine leaves after Plasmopara viticola *infection*

The thesis work was conducted at the Proteomics Laboratory of the Biotechnology Department of the University of Verona, in collaboration with other laboratories: concerning the study on colorectal cancer cells, the Laboratory of Oncology of the IRCCS Foundation “Istituto Nazionale dei Tumori”, Milan, provided all the biological samples, whilst the multivariate analysis of protein profiles was possible thanks to the collaboration with the Department of Environmental and Life Sciences of the University of Eastern Piedmont, Alessandria.

The biochemical and proteomic analysis of *lactobacillus plantarum* wine strain was the result of the collaboration with laboratory of Dr. Zapparoli (Department of Biotechnology of the University of Verona).

Proteomic investigations on the Grapevine leaves infected by *Plasmopara viticola* were performed in collaboration with laboratory of Dr. Polverari (Department of Biotechnology of the University of Verona).

Finally, the identification of proteins for all the proteomic analyses performed were possible thanks to the collaboration with the Proteomics laboratories of Department of Environmental Sciences, Tuscia University, Viterbo, and of International Centre for Genetic Engineering and Biotechnology, Trieste.

The results thus obtained are here discussed and evaluated.

CONTENTS

LIST of ABBREVIATIONS

2-D: two-dimensional

2-DE: two-dimensional electrophoresis

2D-PAGE: two-dimensional polyacrylamide gel electrophoresis

AA: acrylamyde

APC: adenomatosis polyposis coli

BE: backward elimination

CA: carrier ampholytes

CHAPS: 3-[(3-cholamidopropyl)dimethylammonium]-1-propane-sulfonate

CRC: Colorectal cancer

CSF: cerebrospinal fluid

DIGE: differential in-gel electrophoresis

DTT: dithiothreitol

ECL: enhanced chemiluminescence

ESI: electrospray-ionization

GO: gene ontology

HDAC: histone deacetylases

ICAT: isotope-coded affinity tag

IEF: isoelectric focusing

IPG: immobilized pH gradient

IT: ion trap

LAB: lactic acid bacteria

LC: liquid chromatography

LCM: laser capture microdissection

LC/MS-MS: liquid chromatography tandem mass spectrometry

LV: latent variable

MALDI: matrix-assisted-laser-desorption/ionization

MLF: malolactic fermentation

Mr: relative mass

MS: mass spectrometry

NCBI: national center of biotechnology information

OD: optical density

PCA: principal component analysis

PCs: principal components

pI: isoelectric point

PLS-DA: partial least squares – discriminant analysis

PR: pathogenesis-related

PTMs: post-translational modifications

PVDF: polyvinylidene difluoride membrane

PVPP: polyvinylpolypyrrolidone

Q-TOF: quadrupole-time-of-flight

RP-HPLC: reversed phase high performance liquid chromatography

SC: single component

SDS: sodium-dodecyl-sulphate

SDS-PAGE: sodium-dodecyl-sulphate polyacrylamide gel electrophoresis

SELDI: surface enhanced laser desorption ionisation

SIMCA: soft-indipendent model of class analogy

SSP: standard spot number

TA: tannic acid

TBP: tributyl phosphine

TCA: trichloroacetic acid

TOF: time of flight

CHAPTER 1

TOOLS AND APPLICATIONS OF PROTEOMICS

1.1 GENERAL INTRODUCTION

The terms proteomics and proteome were coined by Marc Wilkins and colleagues in the early 1990s, and have then been adopted by the research community at large. The proteome term refers to the full complement of proteins encoded by the genome of an organism [1, 2]. Biochemical studies of protein activity have historically focused on the analyses of single molecular species. The rapid discovery of new gene products, via large-scale genomic initiatives, has necessitated the development of alternative strategies to evaluate protein function. The challenge in recent years has been to develop high-throughput approaches to facilitate systematic protein analysis of biological samples, map functional interactions between proteins on a global scale, and place them in a biological context.

To really understand biological processes, we need to understand how proteins function in and around cells since they are the functioning units. Proteomics, the science that studies the proteome, includes not only the identification and quantification of proteins, but also the determination of their localization, modifications, interactions, activities, and, ultimately, their function. Cells express genes encoding proteins with unique, cell-specific functions in addition to proteins that carry out generic, yet essential functions. Approximately 200 different post-translational modifications (PTMs) have been reported, encompassing a wide variety of reversible and irreversible chemical reactions, which influence the diversity, affinity, function, cellular abundance and transport of proteins. The proteome comprises all of these forms, which makes its analysis incredibly complex, as any protein may be present in a cell, in any form, at any given time. As a consequence, every organism possesses one genome and multiple proteomes. Proteomes are constantly changing in response to environmental stimuli, chemicals and drug treatments, as well as growth and disease processes [3]. Many of these changes hold considerable interest for researchers seeking to understand

complicated disease pathologies, such as cancer. As the messenger RNA abundance in a cell often poorly correlates with the amount of proteins synthesized, the proteome, rather than the transcriptome, holds promise for the identification of novel therapeutic targets. While proteomics initially encompassed just two-dimensional (2-D) gel electrophoresis for protein separation and identification, now it refers to any procedure that characterizes large sets of proteins. The explosive growth of this field is driven by multiple forces: genomics and revelation of more and more new gene products; powerful protein technologies, (such as newly developed mass spectrometry approaches, global yeast two-hybrid techniques, and spin-offs from DNA arrays); and innovative computational tools and methods to process, analyze, and interpret prodigious amounts of data.

1.2 SAMPLE PREPARATION AND PROTEIN EXTRACTION

1.2.1 SOURCE OF SAMPLES

Tissues, cell lines, primary cell cultures and body fluids (such as plasma or cerebrospinal fluid (CSF)) can be used as protein source. Different branches of proteomics are also devoted to plant, bacteria and virus proteins. Anyway, independently from the experimental model, sample preparation is a matter of great importance, especially in comparative proteomics. Thus, not only an appropriate experimental model, but also an optimized protein extraction protocol, have to be carefully considered to obtain reliable results.

1.2.1.1 Animals

When a particular animal model has been established, usually the tissue connected with a particular disease is chosen for detailed analysis. It’s obvious that every tissue has its own characteristics and has to be handled properly for protein extraction. During tissue preparation for proteomic analysis, it is important to diminish its heterogeneity, as much as it is possible. The sample should be pure and relevant. For example, in case of cancer proteome analysis, it should be free of stroma, blood, serum, etc.. Fresh tissue

should be free from connective tissue and fat. Biopsy might be another source of samples for proteomics analysis. It seems to reflect the state of living organism, and sometimes, collected material could be cultured for further experiments. On the contrary, simplification of the sample is one of the greatest advantages of cell culture serving as experimental model, especially in proteomics. This model allows for studying the behaviour of a single type of cell in the absence of the complexity of the entire tissue. This may help to reveal changes in low abundance proteins, which could be impossible in the whole tissue study. Body fluids are another important source providing vital information on the function of living organisms. Body fluids (such as blood, CSF, urine, and saliva) are relatively easily available; thus they are commonly used for clinical diagnosis. Anyway, great difficulties are associated with the broad dynamic range of components present in body fluids [4].

1.2.1.2 Bacteria

Concerning bacterial samples preparation, problems can arise in disrupting cells, due to the presence of thick cell walls and polysaccharide capsule in certain bacterial groups [5]. Some bacteria could simply be lysed by the constituents of the lysis buffer (reducing agents and detergents), but others must be disrupted mechanically (e.g. by sonication). Sometimes, removal of the cell wall by enzymatic digestion is necessary.

1.2.1.3 Plants

A characteristic of plant cell is the presence of cell wall, mostly made up of cellulose and its derivatives. Generally, disruption of cell wall and protein releasing are crucial for analytical success [6]. Another specific feature of plant proteome analysis is the presence of non-proteinaceous contaminants specific of plants (such as polyphenols, lipids, organic acids, terpenes or pigments [7]) that can interfere with separation methods. Therefore, removal procedures by protein precipitation are desirable; for instance, acetone or 10% trichloroacetic acid in acetone are the most frequently used.

1.2.2 METHODS OF CELL DISRUPTION

The first step in sample preparation for proteomic analysis is cell disruption. One of the simplest method is homogenization that helps to break down the initial sample, to finer particles in the nano- to micro-meter scale, so that target molecules, such as proteins, can be released from the tissue for further processing. Ultrasonic homogenizers (also called disintegrators), are based on the piezoelectric effect. Alternatively, pressure homogenizers (such as French press) are an effective system for homogenization of eukaryotic, as well as microbial and plant cells in suspension. In addition, French press is often applied for the preparation of cell membranes for further experiments. Finally, osmotic lysis methods utilize osmotic pressure to destroy cell walls and membranes. Detergents are also efficient for cell disruption and extraction of nuclear and mitochondrial membranes. The most commonly applied detergents are Triton X-100, Tween 80, Nonidet P-40 (NP 40) and saponin.

1.2.3 PROTEIN SOLUBILIZATION

Protein solubilization strongly affects the quality of final results and thus determines the success of the entire experiment. Once isolated, proteins are often insoluble in their native state. To avoid protein modifications, aggregation or precipitation (resulting in the occurrence of artefacts and subsequent protein loss), sample solubilization implicates the use of chaotropes, detergents, reducing agents and protease inhibitors.

Usually, a neutral chaotropic agent, urea, is used at high concentrations ranging from 5 to 9M to effectively disrupt secondary protein structure. As indicated by Rabilloud et al. [8], addition of thiourea to the denaturing solution containing urea, allows for substantial improvement of protein solubility. Addition of thiourea to the sample buffer decreases solubilization of urea. Therefore, when combined with 2M thiourea, urea concentration should not exceed 5–7M [9].

Detergents and amphipathic molecules enable protein extraction and solubilization. It has already been reported the great solubilizing power of zwitterionic detergents [10, 11], in addition it has been proved that non-ionic detergents are also efficient [12, 13]. Nowadays, sulfobetaine 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) is most commonly applied in proteomic studies due to its high solubility and a

relative lack of detergent-induced artefacts. As regarding solubilization of the integral or membrane-associated proteins it has to be improved by using special detergents (e.g. Triton X-114 [14]).

Reducing agents disrupt disulfide bonds between cysteine residues, thus, promote protein unfolding and enable analysis of single protein subunits. Conventionally, phosphines, e.g. tributhylphosphine (TBP) and triscarboxyethylphosphine (TCEP) are used as sulfhydryl reducing agents: they are apolar, therefore don't migrate out of the pH gradient while performing isoelectric focusing. Moreover phosphines do not interact with the alkylating substrates such as 4-vinylpyridine and acrylamide [15]. Thus, reduction and alkylation may be performed in a single step.

Finally, if not inhibited during cell membrane disruption, endogenous proteases are responsible for uncontrolled enzymatic proteins degradation. Such proteolysis may produce artefacts and hence complicate further analysis [16]. In general, addition of specific protease inhibitors, or cocktails of them with a broader activity spectrum, is recommended [17, 18].

1.2.4 REMOVAL OF CONTAMINANTS

Among most common contaminants, salts and lipids naturally occur in body fluids and tissues, and may interfere with electrophoretic separation of proteins [19]. Most often, salt removal is being accomplished via (spin, micro) dialysis [20, 21], ultrafiltration [22, 23], gel filtration, and precipitation [24]. In addition, the use of centrifugal filter device and the sample buffer including CHAPS allows for efficient lipid and salt removal. Alternatively, precipitation in acetone or in trichloroacetic acid (TCA)/acetone removes lipids efficiently. Polysaccharides and nucleic acids are negatively charged polymers which interact with proteins and clog the pores of the polyacrylamide gels causing streaking visible on 2D gels. In order to exclude polysaccharides and nucleic acids from the sample, precipitation may be beneficial (e.g. by complexing them with carrier ampholytes). Moreover in order to remove DNA and RNA, digestion with protease free DNase and RNase is often applied [25]. The presence of phenols observed in plant tissues, may modify proteins through an enzyme-catalyzed oxidative reaction. Oxidation may be prevented during tissue extraction by reducing agents. Furthermore, protein

precipitation with TCA, followed by extraction of phenols with ice-cold acetone or phenol adsorption to polyvinylpolypyrrolidone (PVPP) [26], are advantageous.

Finally, buffers, salts and detergents (included in solubilizing solutions) often tend to interfere with further protein separation steps: they inhibit the in-gel digestion process and interfere with the mass spectrometry analysis. Thus they need to be removed at a proper time of analysis. Protein precipitation followed by resuspension in an adequate sample solution belongs to the most commonly applied procedures enabling removal of most of contaminants. Unfortunately, there is currently no method that would allow precipitating all proteins and, consequently, only precipitated proteins can be further resolubilized. Most commonly, precipitation with TCA, acetone, chloroform/methanol, ammonium sulphate or combinations of the above are being performed.

1.2.5 PROTEIN DEPLETION AND ENRICHMENT METHODS

The proteome and subproteomes of any living cell are highly dynamic and of unknown complexity, making the characterization of proteomes a formidable challenge. While the concentration range of proteins in a given sample may exceed 10 orders of magnitude [27], currently available proteomic approaches are estimated to focus on the 30% most abundant proteins in an extract. Proteomic analyses are, hence, complicated and detection of least abundant proteins is hampered by those molecules present at higher concentration. Plasma, serum and CSF, are protein sources of great importance to biomedicine, diagnostics and therapeutics, containing up to 90% of highly abundant proteins such as albumin, immunoglobulins (IgG and IgA), antitrypsin, transferrin, transthyretin, α1-antitrypsin, hemopexin and haptoglobin. Removal of these proteins may increase detection of other molecules present at lower concentration. Various strategies have been presented for the removal of high-abundant proteins [28], most of which are based on affinity chromatography employing mimetic ligands [29, 30], proteins A and G [31], and antibodies.

On the other side, enrichment methods allow for increasing the concentration of proteins of interest. This statement is really important in the proteomics study, because usually low-abundance proteins carry valuable diagnostic information and are responsible for processes ongoing in cells. In order to increase the likelihood of

detecting less abundant proteins, a variety of protein separation techniques are being applied to reduce the complexity of a sample prior to proteome analysis [32, 33]. The fundamental idea of prefractionation is to isolate the sample into distinct fractions containing restricted numbers of molecules. Preparative liquid electrophoretic methods are common fractionation systems for proteome profiling. The latest technology platforms that have been developed for proteomics include the Rotofor, a multicompartment instrument capable of fractionating proteins according to their pI [34]. Alternatively, common methods of protein enrichment and purification rely on selective precipitation and immunoprecipitation. A novel approach for mining the "unseen proteome" is ProteoMiner^{TM} technology (Bio-Rad), which exploits solid phase combinatorial peptide ligand libraries [34-36]. ProteoMiner^{TM} simultaneously reduces the concentration of high abundance proteins and enriches the concentration of trace proteins without depletion of any specific proteins, in a high throughput manner. Each bead binds a specific binding partner and a wide diversity of ligands is available (typically millions). The population of beads has such diversity that a binding partner should exist for most, if not all, proteins present in a sample. In this system the capacity for high abundance proteins is, in principle, equal to that of low abundance proteins. Thus, when a complex sample is incubated with this bead library, high abundance proteins saturate their binding partner and excess protein is washed away, whilst low abundance proteins are "concentrated" on their specific affinity ligand. In this way the abundance of trace proteins is increased relative to the highly abundant proteins (see Fig. 1). Importantly, in this system, unlike the use of depletion methods, no fraction is discarded and proteins that may be bound to high abundance species like albumin are not lost.

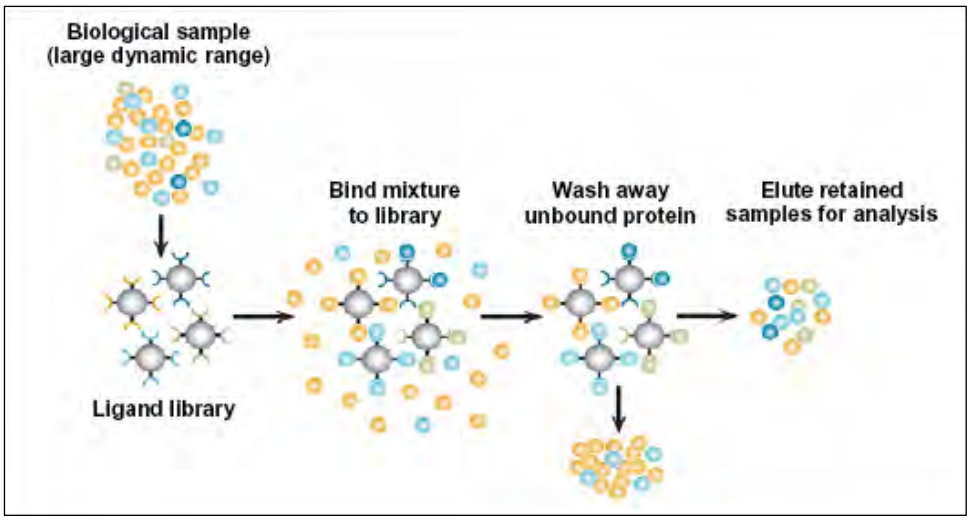

Fig. 1. ProteoMiner™ technology: each bead has a different ligand, with affinity for specific proteins of the sample.

1.3 METHODOLOGIES AND TECHNOLOGIES IN PROTEOME STUDY

Once proteins are purified and ready to be separated, there is a plethora of proteomic techniques to investigate the biological problem the scientist is dealing with. The choice of a given proteomic approach should depend on the type of biological question asked, since each proteomic technology is characterized by specific applications, technical advantages and limitations. In general, proteomic approaches can be used for a) proteome profiling, b) comparative expression analysis of two or more protein samples, c) localization and identification of posttranslational modifications, and d) studying protein–protein interactions. A classical widely used proteome approach relies on two-dimensional electrophoresis coupled to mass spectrometry analysis.

1.3.1 TWO-DIMENSIONAL GEL ELECTROPHORESIS

The traditional two-dimensional gel electrophoresis (2-DE) method was introduced in 1975 by O'Farrell [37] and represents now one of the most commonly applied techniques in proteomics. This method is based on orthogonal separation of proteins according to different physicochemical principles. Proteins in complex mixtures are electrophoretically separated, first according to their isoelectric point (pI) by isoelectric focusing (IEF) and, subsequently, according to their relative molecular mass (Mr) by sodium-dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). Depending

on the gel size and pH gradient, it is possible to resolve simultaneously more than 5000 protein spots [38]. 2-DE is used for large-scale protein separation since its invention for purification, identification as well as quantification of proteins (Fig. 2). Moreover, 2-DE provides information on changes in protein expression, as well as on different protein isoforms and post-translational modifications [39].

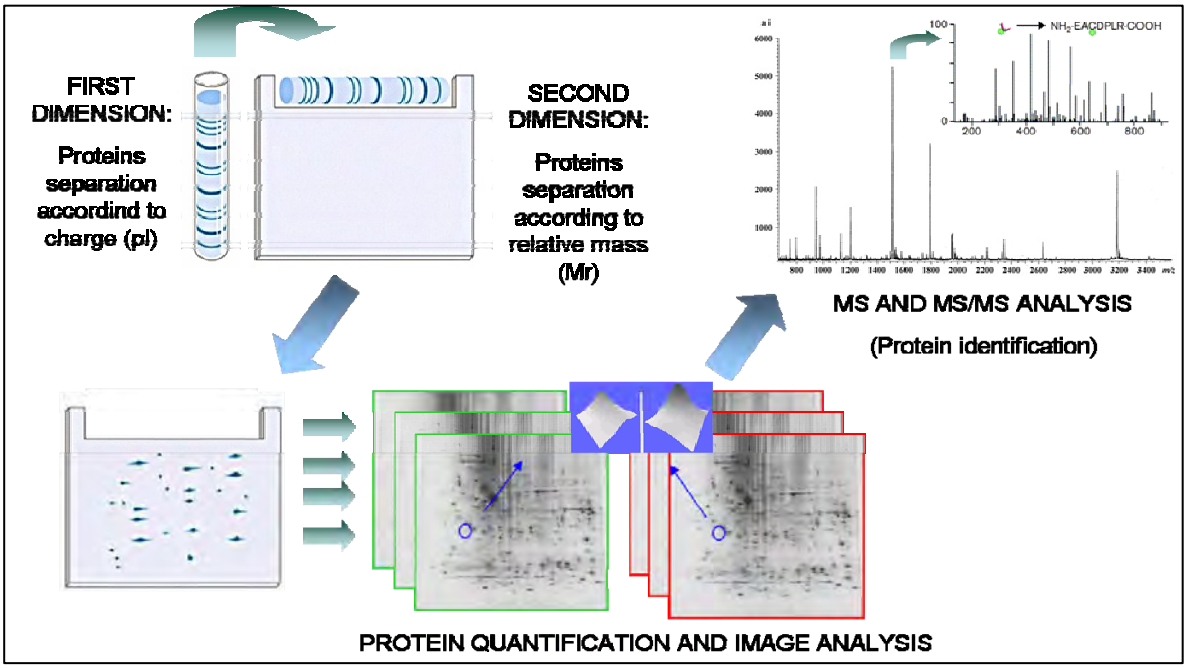

Fig. 2. Overview of two-dimensional gel electrophoresis (2-DE) based strategy.

1.3.1.1 2-DE Methodologies

As mentioned in 1.2.3, the success of any protein separation and purification techniques depends on the protein solubilization method. One traditional sample preparation for 2-DE involves cell or tissue disruption and protein solubilization in presence of high concentrations of urea, reducing agents and detergents. A common 2-DE sample buffer contains 7 M urea, 2 M thiourea, 20 mM Tris, 3% CHAPS. The immobilized pH gradient strip (IPG strip) is then rehydrated with sample, and proteins are separated by IEF.

During IEF, under the influence of an electric field, each protein migrates through the pH gradient until it reaches its pI. Until recently, a mixture of carrier ampholytes was commonly used to create the pH gradient for IEF. However, problems were occurring with these carrier ampholytes-created pH gradients, such as so-called “cathode drifting”

and pH flattening near the anode [40]. In order to circumvent the ampholyte-linked technical difficulties, commercially available immobilized pH gradients [41] are now generally used. The pH gradient in IPG strips is created by acrylamido-buffers (containing weakly acidic or basic buffering groups) admixed to neutral acrylamide monomers and poured gradient-wise via a two-vessel gradient mixer. Covalent immobilization of the buffering groups inside the polyacrylamide gel prevents drifting of the pH gradient under the influence of an electric current, thus increasing the resolution as well as the reproducibility of 2-DE patterns. Moreover, by using IPG strips with a reduced slope of the pH gradient it is possible to significantly increase the resolution of protein separation. Another application of IPG strips is the creation of the so-called zoom-gels, where narrow pH gradients are used allowing for high-resolution separation of spots containing proteins with the same molecular mass but slightly different pI.

After IEF, the IPG strips (with focused proteins) is equilibrated in a SDS-containing solution in order to proceed with the second dimension, where proteins are separated according to their molecular mass by means of SDS-PAGE. Also the resolution of second dimension may be improved by changing the percentage of polyacrylamide and by using pore gradient gels [42]. However, SDS-PAGE separation of high molecular mass proteins as well as very low molecular mass ones is still rather challenging. The applicability of 2-DE analysis generally remains limited to proteins with molecular size between approximately 10 and 150-200 kDa.

1.3.1.2 Protein Detection And Quantification

Protein spots on 2-DE gels can be visualized by a variety of protein staining techniques, each characterized by specific technical aspects, sensitivity, linear range for quantitation, reproducibility and compatibility with mass spectrometric analysis (Table 1). In fact, the major challenge for protein visualization in 2-DE is the compatibility of sensitive protein staining methods with mass spectrometric analysis. Among the most commonly used protein staining methods, Coomassie Blue staining has a sensitivity of up to 30 ng/spot and compatibility with mass spectrometry, while silver staining has an higher sensitivity (1 ng/spot), but was originally incompatible with mass spectrometric analysis, due to the aldehyde fixation steps. Several fluorescent staining methods have been also developed, including Sypro staining and Cy-dyes.

Table 1. Commonly used protein staining methods for 2-DE analysis

Staining Method	Sensitivity	Reversible	MS compatibility	Straightforward protocol	Highly reproducible	Linearity: order of magnitude
Silver	1 ng	-	-	-	±	1
Coomassie	100 ng	+	+	+	+	2
Coomassie R250	50 ng	+	+	+	+	2
Colloidal coomassie	10 ng	+	+	+	+	2
Sypro Ruby	1 ng	+	+	+	+	> 3
Cy-dyes	125 pg	+	+	+	+	4

Sypro Ruby staining [43] allows for a much higher sensitivity, a significantly wider dynamic range, less false-positive staining, and is compatible with mass spectrometric analysis. In addition, Sypro Ruby allows for the detection of lipoproteins, glycoproteins, metalloproteins, calcium-binding proteins, fibrillar proteins and low molecular mass proteins that are very difficult to stain by using other methods. The recently developed Flamingo stain (Bio-Rad) proved to be even more sensitive than Sypro Ruby. Based on Cy-dyes, two-dimensional difference in-gel electrophoresis (DIGE) technology has been developed (Fig. 3). It permits the simultaneous separation of two samples on the same gel by the differential labelling with different cyanine dyes prior to the first dimension [44]. Cyanine dyes carrying an N-hydroxysuccinimidyl ester reactive group covalently bind the ε-amino group of lysine residues in proteins. Staining with these dyes is more sensitive than silver staining, giving a linear response to protein concentrations of up to four orders of magnitude. In addition, no fixation or destaining steps are required during DIGE analysis, reducing protein loss from the gels. Hence, a major advantage of this technique is a significant reduction in inter-gel variability, facilitating spot identification and matching during image analysis, thus increasing the number of analyzable spots.

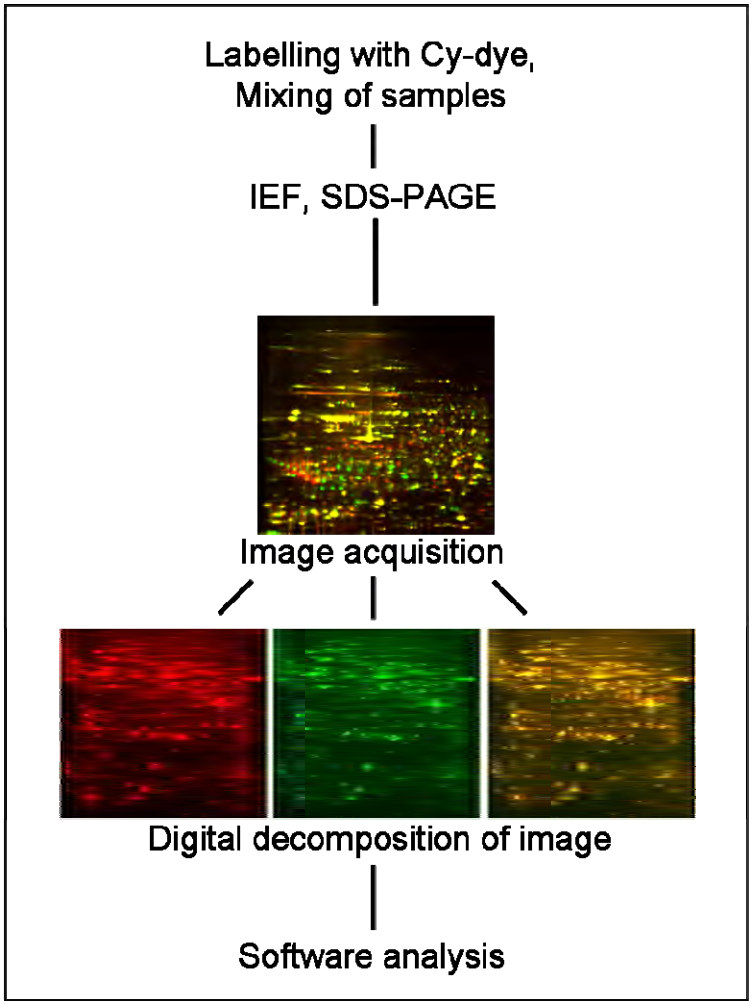

Fig. 3. Flow chart of DIGE analysis.

1.3.1.3 Two-Dimensional Gel Image Analysis

The possibility of detecting protein expression changes (associated with diseases and treatments or finding therapeutic molecular targets) has been, among many other applications, a major incentive to the development of dedicated software systems for 2-DE gel image analysis. In the early 1980s, the first packages started to be delivered to the public and some of them have survived the computational evolution of the last two decades. Among these were PDQuest and ImageMaster 2D Platinum (based on Melanie). Table 2 gives a list of the major available packages. In general, besides basic visualisation properties, the major functions of software systems for 2-DE gel image analysis are (i) detection and quantification of protein spots on the gels, (ii) matching of

Table 1. Major commercialised 2-DE image analysis software

Software	Company	Source website
DeCyder	GE Healthcare	www.gehealthcare.com
Delta2D	Decodon	www.decodon.com
ImageMaster 2D Platinum	GE Healthcare	www.gehealthcare.com
PDQuest	Bio-Rad	www.bio-rad.com
Progenesis (formerly Phoretix)	Nonlinear Dynamics	www.nonlinear.com
Proteomeweaver	Definiens/Bio-Rad	www.definiens.com/www.bio-rad.com

corresponding spots across the gels, and (iii) detection of significant protein expression changes. Any other additional feature, such as data management and database integration, may or may not be included. Moreover, detection algorithms included in the packages generally comprise filtering steps to automatically remove streak artefacts and noise spikes. Spot detection algorithms also produce quantitative information of the protein spots, such as the spot area and OD (maximum intensity value in the area). Matching of gel images is a critical process: it depends on the similarity of the spatial distribution of spots across all the gels, which may vary according to experimental gel running conditions and gel scanning. Some tools propose the initialisation of a few corresponding spots representing the same proteins in different gels: a landmarking step. These landmarks will then be used to warp the gel images and correct possible distortions, and consequently improve the matching quality. Anyway, each tool has its own strengths and weaknesses that vary depending on the gel type and experimental conditions. In the software that deals with DIGE gels, only the spot detection function is adapted. Since the same proteins will be localised in the same x and y coordinates on the gels, the spot detection procedure is the same for the co-detected gels. The matching step is thus straightforward and subject to much less errors. Finally, it should be noted that, despite continuous optimization of algorithms used by these software, a time-consuming visual verification remains necessary for all steps of software analyses of 2-DE patterns.

1.3.1.4 Advantages And Limitations Of 2-DE Based Proteomics

2-DE has the enormous advantage to display simultaneously thousand of proteins on a single 2D map, obtaining a picture of any complete proteome being studied. Furthermore, it permits to analyze polypeptides in their full-length form, thus allowing for characterization of single PTMs by detection with specific antibodies, as well as with

specific staining (e.g. Pro-Q® Diamond phosphoprotein stain). However, while the scientific community is digging more deeply in proteomes, also thanks to fractionation methodology, there are still 2-DE based studies which are “re-discovering similar proteomes in different samples”. This observation creates the impression of scratching the surface: it seems that many studies still reveal similar “housekeeping proteomes” as found in many cell and tissue types, or that the candidate biomarkers for a specific disease are often markers for a generic phenomenon underlying several diseases, such as inflammation [45]. Indeed many proteins are expressed at such low levels that they will escape detection, and other technical problems linked with 2-DE proteomic approach remain, such as co-migration of proteins and loss of insoluble ones. Nevertheless, 2-DE remains a very useful method for displaying and quantification of a majority of proteins in almost any kind of biological samples, and its technologies are robust and increasingly reproducible. Moreover, the sensitivity and applicability of presently available technologies are improving continuously, and for many types of biological questions, specific proteomic approaches are currently available.

1.3.2 MASS SPECTROMETRY TECHNOLOGIES

The invention of mass spectrometers made possible the analysis of proteins and peptides – molecules that have classically been studied individually or in small numbers – in a large-scale and in a high-throughput mode [1]. Remarkable progress has been achieved in the meantime: current mass spectrometric platforms can cover a dynamic range of up to 10^4 [46], and extensive protein and peptide pre-fractionation and sophisticated data acquisition and processing are in place. Today's shotgun proteomic approaches reveal up to some 2000 protein identities and the number of true positive protein identifications within those long lists has enormously increased over the years. This important achievement can be attributed to (a) improved mass spectrometric instrumentation with better sensitivity and specificity, the latter based on superior mass accuracy and resolution (e.g. Orbitrap [47]); and (b) more sophisticated software for peptide sequencing and protein identification.

1.3.2.1 Protein Identification By Mass Spectrometry

In a typical proteomic experiment, proteins of interest are digested with enzymes for identification by mass spectrometric analysis. Trypsin is the most commonly used protease in proteomic analysis and displays good activity with in-gel digestion, as well as with in-solution digestion.

Mass spectrometers consist of an ion source that converts analyte molecules into gas-phase ions, a mass analyzer that separates ionized analytes on the basis of m/z ratio, and a detector that records the number of ions at each m/z value. Regarding ion sources, matrix-assisted laser desorption/ionization (MALDI) [48] and electrospray ionization (ESI) [49] ionization are central in proteomics.

MALDI is generally more salt-tolerant than ESI and it is typically coupled to time-of-flight (TOF) analyzers. For successful MALDI analysis, the key point is the proper choice of matrix and sample deposition method, to achieve the highest possible sensitivity and accuracy. The most popular matrices for proteomic applications include α-cyano-4-hydroxycinnamic acid (CHCA), sinapinic acid (SA) and 2,5-dihydrobenzoic acid (DHB). Particularly, MALDI mass spectrometry, thanks to its speed, sensitivity, accuracy, satisfactory tolerance to impurities and ease of automation, allows for protein identification by the set of measured proteolytic peptide masses. This process is known as peptide mass fingerprinting (PMF) and Mascot is among the most used algorithms in this identification approach [50]. Briefly, the experimental mass profile is matched against those generated in silico from the protein sequences in the database using the same enzyme cleavage sites. Unfortunately, the lack of sequence data makes identification by PMF ambiguous. The recent introduction of LC-MALDI technology offered a sophisticated and efficient mean of sample purification and separation combining the advantages of capillary liquid chromatography (LC) with the sensitivity and accuracy of MALDI-MS; in fact in the newest instruments, peptides eluted from the column are directly spotted onto target plate by robot machines [51].

ESI, instead, is often coupled to ion trap (IT) or quadrupole (Q) analyzers as this combination enables efficient peptide sequencing by induced fragmentation (MS/MS). In an ESI ion source the ionization process occurs under atmospheric pressure and the sample is introduced as liquid, therefore electrospray may be coupled directly to liquid chromatography systems. In particular, reversed-phase liquid chromatography coupled

to tandem mass spectrometry (RP-LC-MS/MS) is a well established and commonly used procedure for identification of the in-gel digested proteins. Furthermore, newly introduced nano HPLC combined with ESI-MS/MS has emerged among the most widely used and the most powerful tools for proteomic research. Indeed, miniaturizing the HPLC separation column inner diameter improves electrospray sensitivity, because the concentration of equally abundant analytes in the LC mobile phase is proportional to the square of the column internal diameter. Moreover, because the sample must be loaded at flow rates of 200 nL/min, the sample volume to be injected onto a conventional nano HPLC system is usually only 1-2 µL. This approach usually requires an additional salt removal step during peptides preparation. In fact, a disposable C18 RP extraction tip, to remove insoluble particles and salts, and to preconcentrate peptides has been introduced [52]. For similar reasons, commercially available capillary chromatography systems include trapping pre-columns, where the sample is purified, desalted and pre-concentrated prior to injection onto the capillary column. In proteomic applications, the positive ion mode is mostly used, in which ionization is based on protonation of the analyte molecules, therefore the sample is often acidified, e.g. with formic or trifluoroacetic (TFA) acids, to facilitate ion formation.

Very recently, a new ionization technique was devised termed desorption electrospray ionization (DESI). Here the charged droplets of solvent are sprayed onto the analyzed object, so that molecules present on its surface are ionized [53]. DESI can be applied to solids, liquids (including complex biological samples) and adsorbed gases. Importantly, DESI apparently does not require sophisticated sample pre-treatment and tissue sections could be directly analyzed.

The mass analyzer is central to mass spectrometric technology, and in the proteomics context, its key parameters are sensitivity, resolution, mass accuracy and ability to produce information-rich fragment mass spectra from peptide ions (tandem mass or MS/MS spectra). There are five basic types of mass analyzers currently used in proteomics [54]: ion trap (IT), time-of-flight (TOF), quadrupole (Q), Fourier transform ion cyclotron resonance (FT-ICR), and the newly developed Orbitrap system [47]. Often, they work as stand-alone mass analyzer, but the current trend points towards hyphenated systems in order to combine the advantages of different analyzers in one mass spectrometer (Fig. 4). In reflector TOF instruments (Fig 4a), the ions are

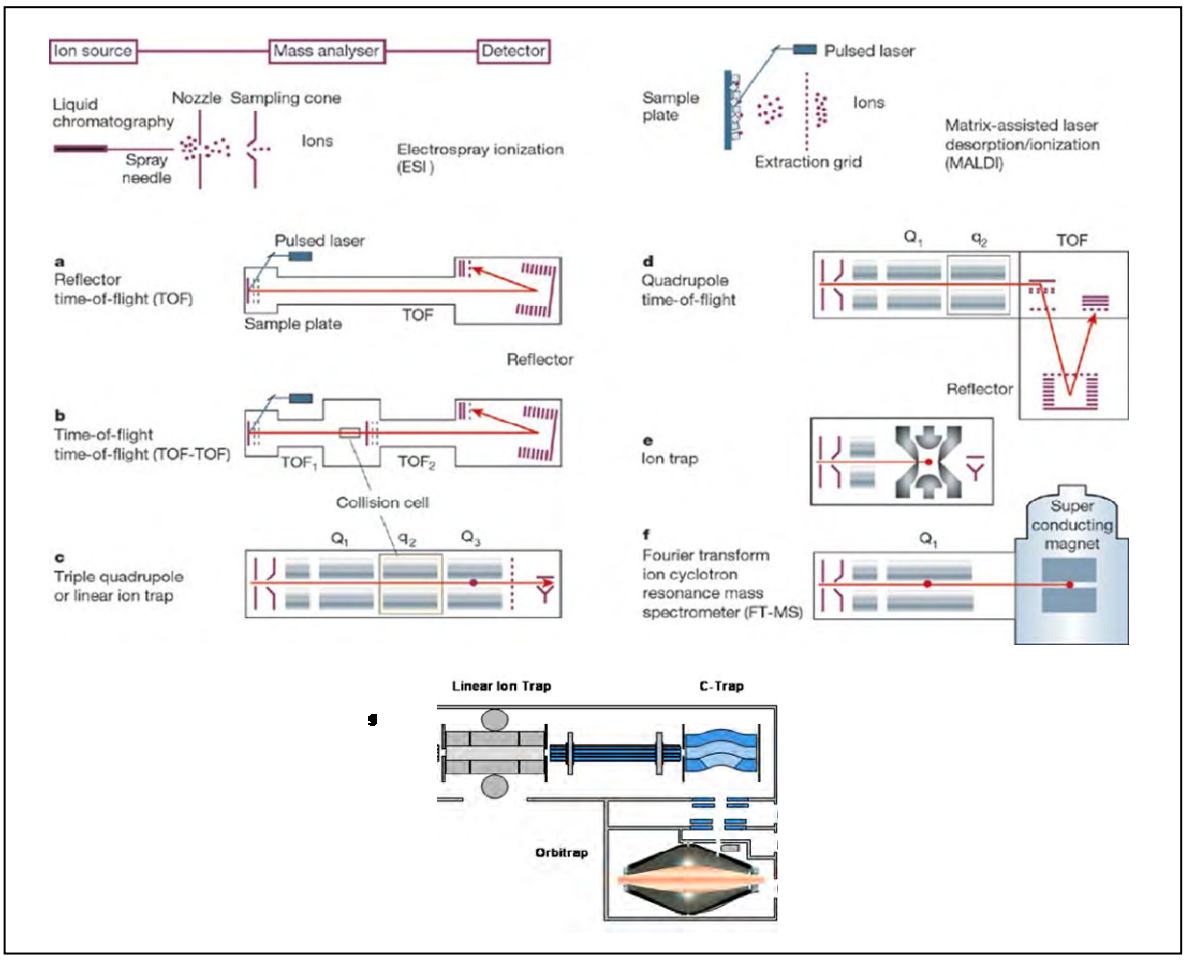

Fig. 4. Mass spectrometers used in proteomic research. The left and right upper panels depict the two ionization technologies, electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI). The different instrumental configurations (a–f) are shown with their typical ion source.

accelerated to high kinetic energy and are separated along a flight tube as a result of their different velocities. The ions are turned around in a reflector, which compensates for slight differences in kinetic energy. The TOF-TOF instrument (Fig. 4b) incorporates a collision cell between two TOF sections. Ions of one mass-to-charge (m/z) ratio are selected in the first TOF section, fragmented in the collision cell, and the masses of the fragments are separated in the second TOF section. Quadrupole mass spectrometers select ions by time-varying electric fields between four rods, which permit a stable trajectory only for ions of a particular desired m/z. In the triple quadrupole analyzer or linear ion trap (LTQ) (Fig. 4c), ions of a particular m/z are selected in a first section (Q1), fragmented in a collision cell (q2), and the fragments separated and scanned out in Q3. The Q-TOF instrument (Fig. 4d) combines the front part of a triple quadruple instrument with a reflector TOF section for measuring the mass of the ions. The (three-

dimensional) ion trap (IT) (Fig. 4e) captures the ions as in the case of the linear ion trap, fragments ions of a particular m/z, and then scans out the fragments to generate the tandem mass spectrum. The FT-ICR instrument (Fig. 4f) also traps the ions, but does so with the help of strong magnetic fields. In the Orbitrap technology (Fig. 4g) ions orbit around a central spindle-like electrode and oscillate harmonically along its axis with a frequency characteristic of their m/z values.

On the basis of this development, a new hybrid mass spectrometer became commercially available very recently: it consists of a LTQ coupled to a IT and the Orbitrap. It combines the robustness, sensitivity, and MS/MS capability of the LTQ with very high mass accuracy and high resolution capabilities of the Orbitrap. The IT-FT-ICR and IT-Orbitrap instruments are especially efficient when combined with fragmentation techniques such as electron capture dissociation (ECD) [55] or electron-transfer dissociation (ETD) [56].

Irrespective of the type of ion source and mass analyzer, MS/MS provides access to sequence data, which allows to more confidently identify peptides. In an MS/MS experiment, daughter ions obtained from fragmentation of a precursor ion produce a unique signature which can be used for database searching. This process is described as peptide fragment fingerprinting (PFF) as opposed to PMF. Identification of proteins using MS/MS data is nowadays performed using three different approaches: (i) peptide sequence tag [57], (ii) cross-correlation method and [58], (iii) probability based matching (typical example is Mascot [50]). All previously described approaches are based on the assumption that a proteomic or translated genomic databases is available to perform the search. *De novo* sequencing approach, instead, uses the MS/MS spectra as the only reference to directly deduce the peptide sequence using specifically developed string similarity search algorithms. They require spectra of higher quality with smaller fragment errors (such as those obtained with FT-ICR or Orbitrap instruments). Currently, PFF is the most common strategy for identifying proteins because of the greater robustness than PMF or *de novo*.

1.3.2.2 MS-Based Proteomics

MS-based proteomics gives the possibility to identify and quantify as many proteins as possible in a single experiment. Nowadays, LC-MS/MS systems have become one of

the preferred methods to conduct large-scale characterization of proteomes. High-performance liquid chromatography (HPLC) is an important separation technique in proteomics. It can easily be coupled to mass spectrometry, which makes it a perfect tool for separation of proteins and peptides directly prior to mass analysis. Liquid chromatography may be used both in top–down and bottom–up proteomics approaches. In the first case, the protein sample is separated and then individual proteins (or simple mixtures) are identified directly by means of tandem mass spectrometry. In the bottom–up approach, the protein, or protein mixture is digested. Single- or multidimensional liquid chromatography coupled to mass spectrometry is then used for separation of peptide mixtures and identification of their compounds. In particular, strong cation exchange (SCX) chromatography is usually adopted as a part of two-dimensional systems, where sample components are first separated on an SCX column and then transferred to the RP column. Chromatographic separation also lowers sample complexity, which makes the analysis more sensitive for low-level components.

While still producing excellent results, “gel-based” quantitative proteomics is going to be superseded by “gel-free” MS-based quantitative proteomics approaches, where quantification is performed using the mass spectrometric data. In both MALDI- and ESI-MS the relationship between the amount of protein present and the measured signal intensity is measured via different approaches. A first solution came with the technique of stable-isotope dilution. This method is based on the fact that pairs of chemically identical molecules (in this case peptide pairs), but with different stable-isotope composition (^{2}H instead of ^{1}H, for example) can be differentiated in a mass spectrometer owing to their mass difference only. Thus the ratio of signal intensities for such peptide pairs should be a direct and accurate measure of the abundance ratio between the two peptides/proteins derived from two different biological conditions. To accomplish this purpose, three main approaches exist today which are: (i) metabolic stable-isotope labelling (for example by amino acids in cell culture, SILAC [59]), (ii) isotope tagging by chemical and (iii) stable-isotope incorporation via enzyme reaction [60]. Regarding the second method, a wide variety of isotopically labelled chemicals has been reported [61]. Isotope-coded affinity tagging (ICAT) [62] was the first approach described in 1999 by Gygi and co-workers. Recently, Gygi and colleagues have described a new method called catch-and-release (CAR) [63]. Several strategies have

been reported that target amines isobaric tag for relative and absolute quantification (iTRAQ) [64]. A clear advantage of all these chemical approaches is the multitude of available functional groups in proteins allowing designing almost any kind of quantitative tag. However, reactions have to be specific, proceed to completion and involve minimal sample handling. Side reactions are problematic, too, as they considerably increase the sample complexity. It has to be noted that, if such experiment is undertaken, one needs to be able to quantify thousands of labelled peptides using automatic tools capable of extracting the intensity for both peptides of a pair and report a protein ratio based on all identified and quantified peptides. Even more important, such tools should be able to process data from different instrument manufacturers and should also be able to accept result input from different search engines (Mascot, SEQUEST, etc.). Several open-source softwares have been developed by scientific laboratories performing large-scale quantitative proteomic experiments and in general each of them has been developed according to supported instruments or database search algorithms.

1.3.3 OTHER PROTEOMIC TECHNOLOGIES

1.3.3.1 SELDI

As a conceptual modification of MALDI-TOF measurements, surface-enhanced laser desorption/ionization (SELDI) technique, that combines chromatographic separation and mass spectral measurement for proteomic profiling and biomarker discovery, is marketed by Bio-Rad. The SELDI chips contain chromatographic coating of selected type (i.e. hydrophobic, ion-exchange, metal-binding, etc.), on which sample components of a given type are captured. Unbound compounds are washed off, thus contaminants are removed and sample complexity is markedly reduced. After application of a proper energy-absorbing matrix, such as CHCA or SA, the proteins bound to the stationary phase are analyzed for MS profiling. The great advantage of SELDI lies in its ability to remove salts and other impurities prior to MS analysis, thanks to which crude samples can be analyzed, such as urine [65], cerebrospinal fluid [66], serum [67], etc.

1.3.3.2 Protein Microarray

Protein microarrays are valuable platforms for both classical and functional proteome analysis and can provide valuable information at the systems-level which is not possible using other techniques. In protein microarrays, antibodies or purified proteins are immobilized on glass slides, which are then used for screening cellular responses to a stimulus (e.g. challenge with plant pathogens) directed against the immobilized proteins and for screening protein-protein interactions [68, 69]. The application of protein microarray technologies in proteomics is now becoming more widespread and many approaches involving antibody and other protein have been used for generating these microarrays. Protein microarrays can be integrated with new techniques such as surface plasmon resonance (SPR) and MS.

1.4 CANCER PROTEOMICS

Cancer represents an enormous medical problem; its huge incidence representing a vast personal and healthcare cost to nations, as cancer remains expensive to diagnose and treat and many of the treatments have toxic side effects. Although the incidence of many cancers (such as breast and prostate cancer), increases as the population ages, other cancers remain characterized by their onset at an early age (e.g. acute lymphoblastic leukemia and neuroblastoma). The risk factors associated with cancer development are incompletely understood; clearly this area involves interplay between environmental effects and genetic predisposition. During the last 30 years, enormous advances have been made with regard to understanding the mechanisms of carcinogenesis. However, these advances have not been paralleled by comparable findings in treatment results. One important factor that has impeded progress in clinical oncology is the complexity of the cancer disease; each tumor type consists of a large number of subtypes that differ with regard to their spectrum of genetic alterations. Each molecular subtype may be associated with a distinct clinical behaviour and treatment response. Routine cancer diagnosis is generally based on microscopical assessments of morphologic alterations of cells and tissues [70]. However, in a substantial number of cases, diagnosis may be ambiguous and the prognosis of disease difficult to determine. Interestingly, some proteins involved in cancer cell behaviour *in vitro* and *in vivo* have

been defined using proteomics techniques. Examples include proteins implicated in melanoma chemoresistance [71] and proteins with a likely role in chemotherapy failure of patients with B-cell chronic leukemia [72].

Cancer proteomics proposes the analysis of molecular pathogenesis of cancer by analysing global protein expression in tumor cells or extracellular fluids of cancer patients. Proteome studies in cancer can be broadly divided into:

1. Tissue culture studies of normal and cancer cells and animal models of cancer;
2. Identification of cancer markers using cell lines and clinical specimens;
3. Protein expression related to cancer cell behaviour and drug treatments.

An important mission of cancer research is the identification of accurate biological markers that can be used for diagnosis, prognosis and treatment of individual tumors. A biological marker (biomarker) is defined as “a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathological processes, or pharmacological responses to a therapeutic intervention” [73]. Immunohistochemistry represents the standard method for marker determinations in routine pathological tissue diagnostics. Nowadays, proteomic strategies are attracting increasing interest for the identification of tissue markers.

High-resolution 2-DE remains a standard tool for cancer expression proteomics. However, running multiple 2-DE gels is usually not feasible when performing expression proteomics of cancer tissues: many clinical samples do not contain sufficient amounts of protein to be analyzed on multiple 2-DE gels. Nevertheless cell cultures may be an advantageous experimental approach to overcome this limit. Moreover, it is possible to identify low-abundance proteins by combining 2-DE with western blotting and antibody cocktails can be used to simultaneously visualize proteins that are of particular interest in a specific project (e.g., signal transduction components) [74]. With regard to searches for novel biomarkers, 2-DE proteomics has the obvious advantage of measuring primarily high-abundance proteins, which are ideal tumor biomarkers since they can be easily measured. Finally, proteomics has a unique potential since it can be used to analyze the state of protein phosphorylation and other modifications on a large scale, and the state of activation of kinase pathways (to predict metastasis and apoptosis sensitivity) [75].

It has to be said that the statistical power of investigations of, for example, 10-20 tumors is extremely low, and most investigators in the prognostic marker field would agree that false positive results may be obtained by statistical fluctuations in materials of <50 tumors. In spite of these limits, interesting results have been obtained, and several independent studies consistently point to the involvement of some markers (such as Hsp27) in malignant transformation *in vivo*.

1.5 MICROBIAL PROTEOMICS

The genomes of most economically important microbial cells are already sequenced and proteomic technologies can be applied during various steps of process development, starting with the selection and optimization of the functions of the industrial strains, application of the knowledge of cell function in response to the changes of production parameters, validation of the downstream processing, and thorough characterization of the final product.

Proteomic methods in food technology and biotechnology, as well as in quality control were at the beginning rather limited. However, this has changed rapidly and proteomics technology is now becoming more practical, and the terms “industrial process proteomics” [76] and “industrial proteomics” [77] have been introduced. Nowadays proteomic technologies can be used in “classical” fermentation industry, for identification of targets for bioprocess improvement [78]. In addition, proteomics is also used for quality control in industrial processes of other food products of animal origin [79].

Yeast fermentation in food processing has been a tradition for thousands of years, and optimization of the fermentation process, purification, productivity, yield and purity, and characterization of the final process are crucial steps in integral process development. For these reasons, baker’s yeast (*Saccharomyces cerevisiae*) was the first eukaryote to have its complete genome sequenced, and it is also the eukaryote with the most deep investigations of the complete proteome.

On the other hand, bacteria are widely used in biotechnology, for production of proteins and chemicals [80], and also for unusual applications like the recovery of ancient frescos [81]. In both the above-mentioned cases, proteomic techniques are used to

follow the production process, such as biosynthesis of recombinant proteins and digestion of hardened glue. Among the most important bacterial organisms, *Escherichia coli* was one of the first sequenced genomes due to its small genome size and prevalent use in laboratory. As a host organism for production of therapeutic proteins, *E. coli* has been for a long time a topic for intensive proteomic analyses [80, 82]. *Lactobacillus* strains are also important industrial microorganisms in dairy and other food fermentations, for production of lactic acid and other bioconversions, and as a source of bactericidal agents. The physical and molecular responses to environmental stress of different *Lactobacillus* species, important for processing of food and beverages, have already been described extensively [83]. An increasingly important topic in microbial research is that concerning “biofilms”. Biofilms are medically important, accounting for over 80% of microbial infections in the body [84]. Proteomic analyses of biofilm forming microorganisms give important information about their behaviour during industrial processes, infection of the host organism, symbiosis, and their defense against antimicrobial agents.

Because of incomplete genome sequencing of some microbes of major industrial importance, analysis of their proteome during the production process is still not possible. Nevertheless, in the near future, proteomics will play, together with genomics and metabolomics, an important role in process development. Possibly, these technologies will be used to enhance productivity of microbial cells or to influence targeted properties of the final product.

1.6 PLANT PROTEOMICS

Plant survival, and hence crop productivity, depends greatly on the ability to adapt, respond, resist, and tolerate variable environmental conditions. Plants have evolved sophisticated mechanisms to cope with a variety of biotic (pathogens, herbivores, parasitic plants) and abiotic (drought, salinity, UV light, poor and polluted soils, *etc.*) stresses. How such environmental stimuli are perceived and trigger the complex defensive and adaptive signalling networks, leading to altered gene expression, protein and metabolic changes, and growth retardation, and how these events result in resistance/tolerance is of major practical interest. The study of the plant–environment

interaction, along with plant organ development, has been conventionally carried out using biochemical, genetic and, more recently, transcriptomic (qRT-PCR, microarrays, SAGE) techniques. Recently, also proteomics has begun to make a further contribution toward the unravelling of mechanisms involved in stress perception and signal transduction. Proteomic studies on plant responses to symbionts, biotic and abiotic stresses, have been published. Concerning the plant material, a number of genotypes, cultivars, transgenic and mutants have been used for differential proteomic experiments, with the intention of identifying protein markers linked to, or that might account for the phenotypic differences observed [85, 86]. Moreover, proteomics has proven to be a very valuable tool for assessing the substantial equivalence and the safety of food and feed derived from transgenic plants [87, 88]. *Arabidopsis thaliana* and rice have been monopolizing proteomic efforts, since the complete genome sequence for both species was publicly available, which simplifying protein identification by MS analysis. For all the species with no significant amounts of published genomic DNA and EST sequences, success in protein identification efforts will be hampered, resulting in low confidence or poor percentage of the proteins identified. Another aspect to consider when working with complex organisms, as in the case of plants, is the fact that plant organs are composed of different cell types, each having its own proteome signature, or responding specifically to certain effectors or environmental stresses, for example, pathogens. The use of laser microdissection is particularly amenable in plant tissues [89] and its application to plant proteomic studies directed at identifying protein forms involved in specific functions, will help to prevent unwanted dilution effects that could mask cell-specific protein changes. Moreover, RuBisCO (ribulose-1,5-bisphosphate carboxylase/oxygenase), the most abundant single protein on earth, accounts for up to 50% of the soluble protein in leaves and this complicates the analysis of minor proteins in green tissue. The problem is exacerbated under physiological (senescence, seed filling) or stressed conditions, where a significant number of proteolysis products of RuBisCO, displaying variable Mr and pI values, has been identified [90]. Analogous to the treatment of albumins in blood serum samples, the development of rapid protocols to eliminate such majority proteins would be beneficial.

Nowadays 2-DE constitutes the dominant protein separation methodology utilized in plant proteomics, but several steps throughout the process, from protein extraction and

separation to spots visualization can be optimized, resulting in improved resolution and increased reproducibility, and allowing for the detection of low-molecular mass food allergens [91] or membrane proteins [92].

For all these reasons, compared to the other applications (cancer and microbial proteomics), plant proteomics, is at very early stage. Nevertheless, its full potential is going to be fully exploited, as the case of the organelle proteome (chloroplast, mitochondria and membranes), where many advances have been already made [93-95].

1.7 BIBLIOGRAPHY

[1] M.R. Wilkins, J.C. Sanchez, A.A. Gooley, R.D. Appel, I. Humphery-Smith, D.F. Hochstrasser, K.L. Williams, Biotechnol. Genet. Eng. Rev. 13 (1996) 19-50.

[2] M.R. Wilkins, J.C. Sanchez, K.L. Williams, D.F. Hochstrasser, Electrophoresis 17 (1996) 830-838.

[3] T. Reynolds, J. Natl. Cancer Inst. 94 (2002) 1664-1666.

[4] G.L. Hortin, S.A. Jortani, J.C. Ritchie, R. Valdes, D.W. Chan, Clin. Chem. 52 (2006) 1218-1222.

[5] P. Cash, Anal Chim. Acta 372 (1998) 121-145.

[6] W.M. Freeman, S.E. Hemby, Neurochem. Res. 29 (2004) 1065-1081.

[7] W. Wang, M. Scali, R. Vignani, A. Spadafora, E. Sensi, S. Mazzuca, M. Cresti, Electrophoresis 24 (2003) 2369-2375.

[8] T. Rabilloud, C. Adessi, A. Giraudel, J. Lunardi, Electrophoresis 18 (1997) 307-316.

[9] B. Herbert, Electrophoresis 20 (1999) 660-663.

[10] M. Chevallet, V. Santoni, A. Poinas, D. Rouquie, A. Fuchs, S. Kieffer, M. Rossignol, J. Lunardi, J. Garin, T. Rabilloud, Electrophoresis 19 (1998) 1901-1909.

[11] G.H. Perdew, S.W. Schaup, D.P. Selivonchick, Anal. Biochem. 135 (1983) 453-455.

[12] S. Luche, V. Santoni, T. Rabilloud, Proteomics 3 (2003) 249-253.

[13] C.M. Taylor, S.E. Pfeiffer, Proteomics 3 (2003) 1303-1312.

[14] R. Polati, A. Castagna, A. Bossi, N. Campostrini, F. Zaninotto, A.M. Timperio, L. Zolla, O. Olivieri, R. Corrocher, D. Girelli, Proteome Science (in press).

[15] B.R. Herbert, M.P. Molloy, A.A. Gooley, B.J. Walsh, W.G. Bryson, K.L. Williams, Electrophoresis 5 (1998) 845-851.

[16] E. Olivieri, B. Herbert, P.G. Righetti, Electrophoresis 22 (2001) 560-565.

[17] A.J. Rai, C.A. Gelfand, B.C. Haywood, D.J. Warunek, J. Yi, M.D. Schuchard, R.J. Mehigh, S.L. Cockrill, G.B. Scott, H. Tammen, P. Schulz-Knappe, D.W. Speicher, F. Vitzthum, B.B. Haab, G. Siest, D.W. Chan, Proteomics 13 (2005) 3262-3277.

[18] D.E. Terry, E. Umstot, D.M. Desiderio, J. Am. Soc. Mass Spectrom. 6 (2004) 784-794.

[19] X. Yuana, D.M. Desiderio, J. Chromatogr. B 815 (2005) 179-189.

[20] T. Manabe, H. Miyamoto, K. Inoue, M. Nakatsu, M. Arai, Electrophoresis 20 (1999) 3677-3683.

[21] C.C. Lai, G.R. Her, Rapid Commun. Mass Spectrom. 14 (2000) 2012-2018.

[22] L. Jiang, L. Hea, M. Fountoulakis, J. Chromatogr. A 1023 (2004) 317-320.

[23] A. Sickmann, W. Dormeyer, S. Wortelkamp, D. Woitalla, W. Kuhn, H.E. Meyer, Electrophoresis 13 (2000) 2721-2728.

[24] R.J. Simpson, Purifying Proteins for Proteomics: A Laboratory Manual, CSHL Press, New York, 2004, p. 393.

[25] R. Kikkawa, M. Fujikawa, T. Yamamoto,Y. Hamada, H. Yamada, I. Horii, J. Toxicol. Sci. 31 (2006) 23-34.

[26] P.K. Koonjul, W.F. Brandt, J.M. Farrant, G.G. Lindsey, Nucleic Acids Res. 27 (1999) 915-916.

[27] N.L. Anderson, N.G. Anderson, Mol. Cell. Proteomics 1 (2002) 845-867.

[28] D. Sitnikov, D. Chan, E. Thibaudeau, M. Pinard, J.M. Hunter, J. Chromatogr. B 832 (2006) 41-46.

[29] S. Kabir, Immunol. Invest. 31 (2002) 263-278.

[30] A. Sato, D.J. Sexton, L.A. Morganelli, E.H. Cohen, Q.L. Wu, G.P. Conley, Z. Streltsova, S.W. Lee, M. Devlin, D.B. DeOliveira, J. Enright, R.B. Kent, C.R. Wescott, T.C. Ransohoff, A.C. Ley, R.C. Ladner, Biotechnol. Prog. 18 (2002) 182-192.

[31] M. Fountoulakis, J.F. Juranville, L. Jiang, D. Avila, D. Roder, P. Jakob, P. Berndt, S. Evers, H. Langen, Amino Acids 27 (2004) 249-259.

[32] M. Fountoulakis, J.F. Juranville, Anal Biochem 313 (2003) 267-282.

[33] J.R. Yates, A. Gilchrist, K.E. Howell, J.J. Bergeron, Nat. Rev. Mol. Cell Biol. 6 (2005) 702-714.

[34] P.G. Righetti, A. Castagna, P. Antonioli, E. Boschetti, Electrophoresis 26 (2005) 297-319.

[35] L. Guerrier, P.G. Righetti, E. Boschetti, Nature Protocols 3 (2008) 883-890.

[36] P.G. Righetti, E. Boschetti, Mass Spectrom. Rev. 27 (2008) 596-608.

[37] P.H. O'Farrell, J. Biol. Chem. 250 (1975) 4007-4021.

[38] A. Goerg, C. Obermaier, G. Bought, A. Harder, B. Scheibe, R. Wildgruber, W. Wiess, Electrophoresis 21 (2000) 1037-1053.

[39] M. Mukherji, Expert Rev. Proteomics 2 (2005) 117-128.

[40] P.G. Righetti, J.W. Drysdale, Ann N.Y. Acad. Sci. 209 (1973) 163-186.

[41] B. Bjellqvist, K. Ek, P.G. Righetti, E. Gianazza, A. Görg, R. Westermeier, W. Postel, J. Biochem. Biophys. Methods 6 (1982) 317-339.

[42] M. Fountoulakis, J.F. Juranville, D. Roder, S. Evers, P. Berndt, H. Langen, Electrophoresis 19 (1998) 1819-1827.

[43] K. Berggren, E. Chernokalskaya, T.H. Steinberg, C. Kemper, M.F. Lopez, Z. Diwu, R.P. Haugland, W.F. Patton, Electrophoresis 21 (2000) 2509-2521.

[44] M. Unlu, M.E. Morgan, J.S. Minden, Electrophoresis 18 (1997) 2071-2077.

[45] J. Petrak, R. Ivanek, O. Toman, R. Cmejla, J. Cmejlova, D. Vyoral, J. Zivny, C.D. Vulpe, Proteomics 8 (2008) 1744-1749.

[46] P. Lescuyer, D. Hochstrasser, T. Rabilloud, J. Proteome Res. 6 (2007) 3371-3376.

[47] A. Makarov, Anal. Chem. 72 (2000) 1156-1162.

[48] M. Karas, F. Hillenkamp, Anal. Chem. 60 (1988) 2299-2301.

[49] J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong, C.M. Whitehouse, Science 246 (1989) 64-71.

[50] D.N. Perkins, D.J. Pappin, D.M. Creasy, J.S. Cottrell, Electrophoresis 20 (1999) 3551-3567.

[51] K.W. Ro, J. Liu, D.R. Knapp, J. Chromatogr. A 1111 (2006) 40-47.

[52] J. Rappsilber, Y. Ishihama, M. Mann, Anal. Chem. 75 (2003) 663-670.

[53] Z. Takats, J.M. Wiseman, B. Gologan, R.G. Cooks, Science 306 (2004) 471-473.

[54] R. Aebersold, M. Mann, Nature 422 (2003) 198-207.

[55] R.A. Zubarev, D.M. Horn, E.K. Fridriksson, N.L. Kelleher, N.A. Kruger, M.A. Lewis, B.K. Carpenter, F.W. McLafferty, Anal. Chem. 72 (2000) 563-573.

[56] J.E. Syka, J.J. Coon, M.J. Schroeder, J. Shabanowitz, D.F. Hunt, Proc. Natl. Acad. Sci. USA 101 (2004) 9528-9533.

[57] M. Mann, M. Wilm, Anal. Chem. 66 (1994) 4390-4399.

[58] J.K. Eng, A.L. McCormack, J.R. Yates, J. Am. Soc. Mass Spectroms. 5 (1994) 976-989.

[59] S.E. Ong, B. Blagoev, I. Kratchmarova, D.B. Kristensen, H. Steen, A. Pandey, M. Mann, Mol. Cell Proteomics 1 (2002) 376-386.

[60] O.A. Mirgorodskaya, Y.P. Kozmin, M.I. Titov, R. Korner, C.P. Sonksen, P. Roepstorff, Rapid Commun. Mass Spectrom. 14 (2000) 1226-1232.

[61] A. Panchaud, M. Affolter, P. Moreillon, M. Kussmann, J. of Proteomics 71 (2008) 19-33.

[62] S.P. Gygi, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, R. Aebersold, Nat. Biotechnol. 17 (1999) 994-999.

[63] C.A. Gartner, J.E. Elias, C.E. Bakalarski, S.P. Gygi, J. Proteome Res. 6 (2007) 1482-1491.

[64] P.L. Ross, Y.N. Huang, J.N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlet-Jones, F. He, A. Jacobson, D.J. Pappin, Mol. Cell Proteomics 3 (2004) 1154-1169.

[65] M.T. Nguyen, G.F. Ross, C.L. Dent, P. Devarajan, Am. J. Nephrol. 25 (2005) 318-326.

[66] S. Ranganathan, E. Williams, P. Ganchev, V. Gopalakrishnan, D. Lacomis, L. Urbinelli, K. Newhall, M.E. Cudkowicz, R.H.J. Brown, J.R. Bowser, J. Neurochem. 95 (2005) 1461-1471.

[67] D.W. Ho, Z.F. Yang, B.Y. Wong, D.L. Kwong, J.S. Sham, W.I. Wei, A.P. Yuen, Cancer 107 (2006) 99-107.

[68] X. Duburcq, C. Olivier, F. Malingue, R. Desmet, A. Bouzid, F. Zhou, C. Auriault, H. Gras-Masse, O. Melnyk, Bioconjug. Chem. 15 (2004) 307-316.

[69] J. LaBaer, N. Ramachandran, Curr. Opin. Chem. Biol. 9 (2005) 14-19.

[70] R. Fox, M. Hull, Ann. N.Y. Acad. Sci. 687 (1993) 217-223.

[71] M. A. Kern, H. Helmbach, M. Artuc, D. Karmann, K. Jurgovsky, D. Schadendorf, Anticancer Res. 17 (1997) 4359-4370.

[72] T. Voss, H. Ahorn, P. Haberl, H. Döhner, K. Wilgenbus, Int. J. Cancer 9 (2001) 180-186.

[73] B. D. W. Group, Clin. Pharmacol. Ther. 69 (2001) 89-95.

[74] J. C. Sanchez, P. Wirth, S. Jaccoud, R. D. Appel, C. Sarto, M. R. Wilkins, D. F. Hochstrasser, Electrophoresis 18 (1997) 638-641.

[75] D. Cecconi, A. Zamò, E. Bianchi, A. Parisi, S. Barbi, A. Milli, S. Rinalducci, A. Rosenwald, E. Hartmann, L. Zolla, M. Chilosi, Proteomics 8 (2008) 4495-4506.

[76] A. Incamps, F. Hély-Joly, P. Chagvardieff, J.C. Rambourg, A. Dedieu, E. Linares, E. Quéméneur, Biotechnol. Bioeng. 91 (2005) 447-459.

[77] D. Josic, M. K. Brown, F. Huang, Y.P. Lim, M. Rucevic, J.G. Clifton, D.C. Hixson, Proteomics 6 (2006) 2874-2885.

[78] W. Wang, J. Sun, M. Hartlep, W.D. Deckwer, A.P. Zeng, Biotechnol. Bioeng. 83 (2003) 525-536.

[79] C. Piñeiro, J. Barros-Velázquez, J. Velázquez, A. Figueras, J. M. Gallardo, J. Proteome Res. 2 (2003) 127-135.

[80] P.S. Lee, K.H. Lee, Biotechnol. Bioeng. 84 (2003) 801-814.

[81] P. Antonioli, G. Zapparoli, P. Abbruscato, C. Sorlini, G. Ranalli, P.G. Righetti, Proteomics 5 (2005) 2453-2459.

[82] N. Zhang, R. Chen, N. Young, D. Wishart, P. Winter, J.H. Weiner, L. Li, Proteomics 7 (2007) 484-493.

[83] M. De Angelis, M. Gobbeti, Proteomics 4 (2004) 106-122.

[84] D. Davies, Nat. Rev. Drug Discov. 2 (2003) 114-122.

[85] C. Finnie, K. S. Bak-Jensen, S. Laugesen, P. Roepstorff, B. Svensson, Plant Sci. 170 (2006) 808-821.

[86] Y. Liu, T. Lamkemeyer, A. Jakob, G. Mi, F. Zhang, A. Nordheim, F. Hochholdinger, Proteomics 6 (2006) 4300-4308.

[87] M.C. Ruebelt, N.K. Leimgruber, M. Lipp, T. L. Reynolds, M.A. Nemeth, J.D. Astwood, K.H. Engel, K.D. Jany, J. Agric. Food Chem. 54 (2006) 2154-2161.

[88] M.C. Ruebelt, M. Lipp, T.L. Reynolds, J.D. Astwood, K.H. Engel, K.D. Jany, J. Agric. Food Chem. 54 (2006) 2162-2168.

[89] T. Nelson, S. L. Tausta, N. Gandotra, T. Liu, Annu. Rev. Plant Biol. 57 (2006) 181-201.

[90] S. Schiltz, K. Gallardo, M. Huart, L. Negroni, N. Sommerer, J. Burstin, Plant Physiol. 135 (2004) 2241-2260.

[91] E. Boschetti, L.V. Bindschedler, C. Tang, E. Fasoli, P.G. Righetti, J. Chromatogr. A 1216 (2009) 1215-1222.

[92] Y. Kashino, T. Harayama, H.B. Pakrasi, K. Satoh, J. Chromatogr. B 849 (2007) 282-292.

[93] Millar, A. H., Whelan, J., Small, I.,. Curr. Opin. Plant Biol. 2006, 9, 610–615.

[94] D. Macherel, A. Benamar, M. H. Avelange-Macherel, D. Tolleter, Physiol. Plantarum 129 (2007) 233-241.

[95] S. Komatsu, H. Konishi, M. Hashimoto, J. Exp. Bot. 1 (2007) 103-112.

CHAPTER 2

PROTEOMIC ANALYSIS OF COLORECTAL CANCER

2.1 COLORECTAL CANCER

Colorectal cancer (CRC) is one of the most common malignant tumours, accounting for at least 1,000,000 new cases worldwide and leading to more than 500,000 deaths each year. About half of the patients with colorectal cancer develop distant metastases during the course of their disease. Overall CRC incidence is of approximately 5% in the general population [1], and CRC risk is largely determined by environmental factors. Among them, diet unquestionably represents the main exogenous factor in the aetiology of colon cancer [2]. Relative to non-dietary factors, a fraction of bowel cancer might be attributable to smoking [3], whereas several inflammatory conditions of the gastrointestinal tract (such as inflammatory bowel disease and Chron's disease) are known to predispose the individual to CRC [4]. Notwithstanding the epidemiological evidence, very little is known on the molecular and cellular mechanisms underlying protective or predisposing effects modulated by diet, lifestyle and other non-genetic factors on CRC risk. In contrast to environmental modifiers, the role of the main genetic factors involved in establishing and modulating predisposition to colorectal cancer has been elucidated during the last 10 years. These findings served as a paradigm for the molecular basis of tumour initiation and progression towards malignancy in a broad spectrum of tissues and organs. It is generally accepted that patients with an increased familial colorectal cancer risk account for approximately 20-25% of all CRC cases [5]. This significant proportion of CRC cases includes both high penetrance hereditary syndromes, such as familial adenomatous polyposis (FAP, caused by germline APC mutations), MYH-associated polyposis (MAP, from MYH gene), hereditary non-polyposis colorectal cancer (HNPCC, caused by the mismatch repair genes MSH2, MLH1, MSH6, and PMS2), Peutz Jegher syndrome (PJS, associated to STK11 gene), and juvenile polyposis syndrome (JPS, due to mutations in SMAD4, BMPR1A, and ENG). Although hereditary CRC syndromes have been instrumental in the elucidation of

the genetic basis of CRC, their relative contribution to the total CRC burden is low, probably not more than ~5% of the cases, thus indicating that the vast majority (15–20%) of the genetic factors predisposing to CRC is yet to be discovered. Also, the genetic basis of familial CRC is likely to be polygenic and the identification of individual modifying genes has proven to be difficult. Alternatively, the relatively high incidence of sporadic CRC cases in the general population may occasionally result in familial clustering by chance.

2.1.1 MOLECULAR BASIS OF COLORECTAL CANCER

Up to now, all colon cancer cases, sporadic or hereditary, are thought to arise through two different pathways. In 1990, a model pathway for the transition of normal epithelia cells to carcinoma cells [6] was published (Fig. 1). In this model, colorectal cancer was hypothesized to occur by a multi-step process in which four genes are mutated in a specific order.

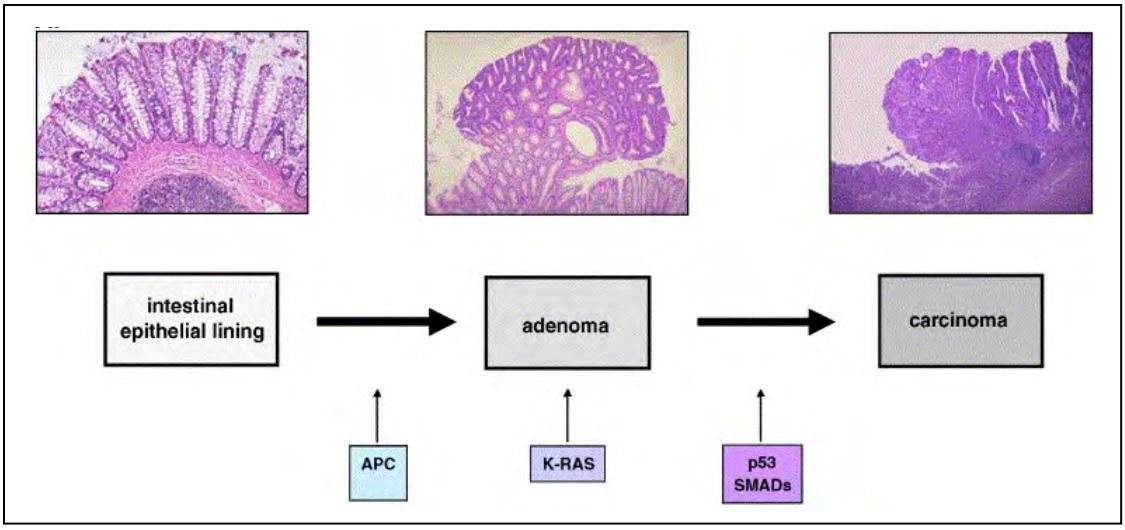

Fig. 1. The adenoma-carcinoma sequence of colorectal cancer and the associated molecular events.

The first occurring step is the loss of function of the APC (adenomatosis polyposis coli) gene, which is a tumour suppressor gene considered a gatekeeper of colorectal cancer. About 25% of APC gene mutations are *de novo* and at the present, over 800 mutations have been characterized. The vast majority of these mutations are unique to a family and most of them cause protein product truncation. The protein product of the APC gene plays a critical role in several cellular processes such as cell division, adhesion, and migration. In particular, this protein ensures that the chromosome number produced

through cell division is correct and it is involved in chromosomal stability. Moreover, the APC protein appears to be involved in cell apoptosis and may decrease cell proliferation. The APC protein accomplishes these tasks mainly through association with other proteins, especially those that are involved in cell attachment and signalling. Researches have indicated that when a truncated version of the protein is made, this results in high levels of free cytosolic β-catenin, a protein which the APC protein usually binds to. This free β-catenin migrates to the nucleus and may bind to transcription factors for oncogenes, resulting in increased cell proliferation or decreased apoptosis, particularly of the mucosal cells of the colon [7]. After the loss of function of the APC gene, activation of the K-ras oncogene is thought to occur. This is followed by loss of gene function on chromosome 18q and inactivation of p53, leading to carcinoma formation [8]. Apart from the above signal transduction pathways, an important class of CRC-involved genes encode for proteins whose main function is of ensuring the maintenance of genome integrity both at the nucleotide and the gross chromosomal levels. In fact CRC is also characterized by chromosomal instability that leads to specific genes or entire portions of chromosomes being deleted or lost, resulting in aneuploidy (abnormal amounts of DNA per cell) [9]. Collectively, these pathways account for approximately 85% of all sporadic colorectal cancer cases and all cases of colorectal cancer associated with familial adenomatous polyposis (FAP).

Subsequent studies have lead to modifications in this model and it is now estimated that at least 7 genes are involved in the pathobiology of CRC [10]. More recently, it was realized that some colorectal cancers occur via a different pathway, characterized by microsatellite instability due to mismatch repair gene mutations. Chromosomal instability and aneuploidy are not involved in this pathway. This pathway has been called the “Mutator Pathway” and it accounts for approximately 15% of all sporadic cancer cases and most cases of hereditary non-polyposis colorectal cancer (HNPCC) [11, 12].

Given the molecular and genetic heterogeneity of colorectal cancer, the completion of the human genome sequence and the advent of genomics and proteomics offer new tools and opportunities for the improved classification of hereditary, familial and sporadic CRC cases.

2.2 SUBCELLULAR PROTEOMICS

One of the main challenges of proteome analysis is the detection of low copy number proteins. There is no amplification step for proteins, analogous to the polymerase chain reaction method for amplifying DNA or RNA. This means that, without any fractionation and enrichment of biological samples, high-abundance proteins hinder low-abundance proteins detection. Most regulatory proteins such as kinases, phosphatases, or GTPases are present at low copy numbers and they are very specific in subcellular localization. In addition, because of the complexity of eukaryotic cells, a single-step characterization of an entire proteome seems, at least presently, rather unfeasible. Indeed, thousands of PTMs, along with the still-underestimated rate of alternative splicing, give rise to a human proteome size that is likely to be significantly larger than the number of estimated genes [13]. Thus, because of the still limited resolution power of proteomic technologies, upstream fractionation steps are needed. Therefore, proteomic research has become increasingly aware of techniques to analyze subcellular proteomes of reduced complexity [14]. Protein patterns of subcellular fractions can be mapped and characterized by high-resolution 2D gel analysis and it was already shown that such an approach may be extremely successful [15]. Organelles and subcellular fractions analysis offers many advantages for proteomics research. Above all, reduced sample complexity makes possible to display an entire organelle proteome on a single 2D gel (or even to analyze all the proteins by mass spectrometric gel-free shot-gun procedure). In addition, identified proteins may immediately be linked to a functional context, because they were purified together with an organelle or subcellular fraction. Thus, the macromolecular architecture of eukaryotic cells can be seen as an advantage for the study of subcellular proteomes, but it is required to establish subcellular fractionation strategies for the individual compartments. Two compartments are of special interest in the light of biomarker identification by proteomic methods: the cell nucleus and the cellular membranes. In particular, the nucleus is the repository for the cell genetic information, in the form of the chromosomes, and it is the site of gene expression and regulation. Therefore identifying nuclear proteins is important to understand mechanisms involved in gene expression control, as well as for providing clues about the molecular function of novel proteins.

Moreover, because of physiological processes occurring during apoptosis, such as DNA degradation and nuclear proteins cleavage [16, 17], nuclear proteome analysis is of great interest in cancer research. In fact, a set of observed proteins possibly reflects very fundamental networks for deregulation of cellular function leading to tumour development.

2.3 A PROTEOMIC APPROACH FOR EVALUATING THE CELL RESPONSE TO A NOVEL HISTONE DEACETYLASE INHIBITOR IN COLON CANCER CELL

2.3.1 INTRODUCTION

The acetylation state of several cellular proteins (including histones and transcription factors) is controlled by the action of histone acetyl transferases and histone deacetylases (HDAC) [18]. The post-translational modification of histones has emerged as a central mechanism of transcriptional control, because hypoacetylation results in transcriptional repression as a consequence of tight packing of DNA into the nucleosome. Modulation of histone acetylation and chromatin relaxation affect diverse cellular functions including control of cell growth, differentiation and apoptosis [19]. In addition to modulation of histone acetylation, HDACs are also involved in the lysine acetylation state of non-histone proteins implicated in regulatory processes (e.g., transcriptional factors and tubulin) [18]. Epigenetic inactivation of gene expression is a general phenomenon associated with malignant transformation [20]. Indeed, the expression of several regulatory genes, including tumour suppressor genes, differentiation genes and DNA repair genes, may be repressed during tumour transformation and progression. For these reasons, modulation of epigenetic gene repression has been proposed as an attractive approach to control tumour growth, and HDAC, are recognized as potential targets of this approach [21, 22]. Class I and class II HDAC, are zinc-dependent metalloenzymes that catalyze the hydrolysis of acetylated lysine residues. Most HDAC inhibitors have a zinc binding group, such as hydroxamic acid. Therefore, these HDAC inhibitors are not isoenzyme-specific and could be

considered as pan-HDAC inhibitors. In a previous experiment the effects of the hydroxamic acid-based HDAC inhibitor trichostatin A (TSA) on pancreatic exocrine [23, 24] and endocrine [25] cancer cell lines have been explored, showing that drug-induced inhibition of proliferation was associated with cell cycle arrest and apoptosis and a profound change in the global proteomic profile. In spite of much effort, the design of selective inhibitors remains difficult and only few compounds exhibited some degree of specificity [26, 27]. It remains to be defined if isoenzyme-specific HDAC inhibitors provide advantages over aspecific pan-HDAC inhibitors. Recently, in the laboratory of Prof. Zunino at the "Istituto Nazionale dei Tumori", Milan, a novel series of HDAC inhibitors containing hydroxamic acid has been found to exhibit a broad-spectrum inhibition profile characterized by a marked effect on acetylation of histone and non-histone proteins. RC307, a representative compound of this series, was able to induce apoptosis in ovarian carcinoma cells. In the human colon carcinoma cells HCT116, which is known to be a cell system responsive to HDAC inhibitors [28], RC307 exhibited as anti-proliferative effect that was associated with G2-phase accumulation but only a moderate induction of apoptosis. Based on these observations, the present study (performed in collaboration with Prof. Zunino group and Prof. Righetti, Department of Chemistry, Materials and Chemical Engineering "Giulio Natta", Polytechnic of Milan) was designed to investigate the pattern of protein expression in tumour cells following treatment with anti-proliferative concentration of RC307. In addition, proteomic findings of this work may elucidate the critical proteins expression alterations responsible for the activity of this novel HDAC inhibitor. Understanding the mechanism of action of RC307 will finally also allow an improvement in knowledge of the epigenetic events involved in colon cancer development.

2.3.2 MATERIALS AND METHODS

2.3.2.1 Cell Culture And Drugs

The human colon cancer HCT116 cell line was grown and maintained as monolayers in RPMI-1640 medium supplemented with 10% fetal calf serum (Invitrogen, Carlsbad, CA). RC307 was dissolved in dimethylsulfoxide and diluted in water.

2.3.2.2 Growth-Inhibition Assays

Cell sensitivity to the HDAC inhibitor RC307 was assessed by growth-inhibition assay. Exponentially growing cells were harvested, seeded into 6-well plates and 24 h later cells were exposed to different concentrations of RC307 (range: 0.1-3 μM) or to solvent for 72 h. At the end of treatment, culture medium was removed and adherent cells were harvested using trypsin and counted with a cell counter (Coulter Electronics, Luton, UK). IC50 is defined as the concentration causing a 50% inhibition of cell growth as compared with control.

2.3.2.3 Apoptosis And Cell Cycle Analysis

Exponentially growing cells were seeded in 75 cm^2 flasks and 24 h later, they were exposed to a drug concentration corresponding to IC_{80} (2.5 μM) for 24, 48 or 72 h. Under these conditions the concentration of the drug solvent (0.005%) had no effects. At the end of treatment, floating and adherent cells were harvested for detection of apoptotic cells or cell cycle analysis. Apoptosis was evaluated by TUNEL (Terminal deoxynucleotidyl transferase dUTP Nick End Labelling) assay (Roche, Mannheim, Germany). After harvesting the cells were fixed in para-formaldehyde, permeabilized in a solution of 0.1% Triton X-100 in 0.1% sodium citrate, and then incubated in the TUNEL reaction for 1 h. After washing, samples were analysed by flow cytometry using Cell Quest software (Becton Dickinson, Mountain View, CA). For cell cycle analysis, cells were fixed and stained with a propidium iodide (PI)-containing solution (30 μg/ml PI, 66 U/ml RNase A in PBS). The cell cycle perturbations were measured by using a flow cytometer (Becton Dickinson, Franklin Lakes, NJ). Samples were analyzed for DNA content and cell cycle distributions were calculated using Modfit software (Becton Dickinson).

2.3.2.4 Total Lysates And Nuclear Protein Extraction

Exponentially growing cells were seeded in 150 cm^2 flasks and, 24 h later, they were exposed to the HDAC inhibitor for 24 h (IC_{80}: 2.5 μM). Total lysates protein extraction from 10^8 untreated and RC307-treated cells was performed in a 2-D solubilising/lysing solution: 7 M urea (Sigma, Sigma-Aldrich Corporation, St. Louis, MO, USA), 2 M thiourea (Sigma), 3% CHAPS (Sigma), 20 mM Tris (Sigma), 1% pH 3-10 Ampholine

(Fluka, Buchs SG Switzerland) and 1x protease inhibitor cocktail tablet (Complete, Mini; Roche, Basel, Switzerland). The samples were then sonicated 5x30 s on ice with 1 min rest in between times and the sonicates were centrifuged for 10 min at 10,000xg at 4°C to remove the nucleic acids complexed with ampholytes. Concerning the nuclear proteins, the extraction from 10^9 untreated and RC307-treated cells was obtained with the CelLytic NuCLEAR Extraction Kit (Sigma) following manufacturer's instructions. Then 1% pH 3-10 Ampholine was added and samples were centrifugated as above to remove the nucleic acids. Both total lysates and nuclear extracts were incubated with 5 mM tributyl phosphine and 20 mM acrylamide for 60 minutes at room temperature to reduce protein disulphide bonds and alkylate the cysteine thiolic groups. The reaction was blocked by the addition of 10 mM DTT (Sigma) and the samples were collected and stored at -80°C. Protein concentration was evaluated with DC Protein assay (Bio-Rad, Labs., Hercules, CA, USA) based on the Lowry method.

2.3.2.5 Two-Dimensional Gel Electrophoresis

Total protein extracts were analyzed by 2D PAGE on 18x20 cm slab gels by using 17 cm immobilized 3-10 pH gradient (IPG, Bio-Rad) and 8-18% polyacrylamide gradient. The IPG strips were rehydrated for 8 h with 450 µL of 2-D solubilizing solution (7 M urea, 2 M thiourea, 3% CHAPS and 20 mM Tris) containing 1.5 mg/mL of total protein. Isoelectric focusing (IEF) was carried out with a Protean IEF Cell (Bio-Rad), with a low initial linear voltage (from 0 V to 1000 V over 12 hours) and then by applying a final rapid voltage ramp up to 10000 V with a limiting current of 50 µA/strip. The total product time x voltage applied was 70,000 Vh for each strip and the temperature was set at 20°C. For the second dimension, the IPGs strips were equilibrated for 26 min by rocking in a solution of 6 M urea, 2% SDS, 20% glycerol, 375 mM Tris-HCl, pH 8.8. The IPG strips were then laid on gradient SDS-PAGE with 0.8% agarose in Tris/glycine/SDS running buffer (192 mM glycine, 0.1% SDS and Tris to pH 8.3) containing traces of bromophenol-blue tracking dye. The second dimension was performed in a Protean Plus Dodeca cell (Bio-Rad) with Tris/glycine/SDS running buffer. The electrophoresis was conducted with continuous cooling (at 18°C) and mixing by setting a current of 40 mA for each gel for 3 min, then 2 mA/gel for 1 h, and 20 mA/gel until the tracking dye, bromophenol-blue, reached the anodic end of the gels. The protein zones were finally

revealed with Sypro Ruby stain (Bio-Rad). Gels were incubated in a fixing solution containing 40% ethanol and 10% acetic acid for 30 min followed by overnight staining in a ready-to-use Sypro Ruby solution. Destaining was performed in 10% methanol and 7% acetic acid for 1 h, followed by a rinse of at least 3 h in pure water.

2.3.2.6 Image Analysis

The image analysis of the 2D gels replicates was performed by PDQuest software (Bio-Rad), version 7.3. Each gel was analysed for spot detection, background subtraction and protein spot OD intensity quantification. The gel image showing the higher number of spots and the best protein pattern was chosen as a reference template, and the spots in a standard gel were then matched across all gels. Spot quantity values were normalised in each gel by dividing the raw quantity of each spot by the total quantity of all the spots included in the standard gel. Two distinct differential analyses were performed, one for total lysates and one for nuclear extracts. In both the experiments gels were divided in two separated groups (control and RC307-treated samples) and, for each protein spot, the average spot quantity value and its variance coefficient in each group were determined. A Student's t-test was performed in order to compare the two groups and identify sets of proteins that showed a statistically significant difference with a confidence level of 0.05.

2.3.2.7 In-Gel Digestion

Spots were carefully cut out from 2-D Sypro Ruby stained gels and subjected to in-gel trypsin digestion. Briefly, spots were destained (1x15 min 300 µl wash in 100 mM NH4HCO3; 1x15 min 300 µl wash in 50% 100 mM NH_4HCO_3 (v/v), 50% acetonitrile; 1x5 min wash in 100% acetonitrile), and dried at 37°C. The gel pieces were then swollen in 10 µl of a digestion buffer containing 100 mM NH_4HCO_3 and 20 ng/µl of trypsin (modified porcine trypsin, sequencing grade, Promega, Madison, WI). After 10 min, 40 µl of 100 mM NH_4HCO_3 were added to the gel pieces and digestion allowed to proceed at 37°C. The supernatants were collected and peptides were extracted in an ultrasonic bath for 10 min (twice 50 µl 50% acetonitrile, 100% H_2O with 1% formic acid v/v; once 50 µl of acetonitrile). All the supernatants coming from the different steps were

collected in the same tube. Tryptic peptides were dried by vacuum centrifugation and redissolved in 20 µl 0.1% formic acid in water.

2.3.2.8 Peptide Sequencing By MALDI-TOF/TOF And Nano RP-HPLC-ESI MS/MS

Tryptic peptides for MALDI TOF/TOF analysis were prepared by diluting 1 µl of peptide solution with 1 µl of a saturated solution of CHCA (10 mg/ml) containing 0.1% TFA and 50% acetonitrile. MS analysis was conducted with a 4800 MALDI TOF/TOF mass spectrometer (Applied Biosystems, Framingham, MA, USA) in positive-ion reflectron mode. For MS analysis, data were acquired automatically after 600 laser shots with a fixed laser intensity of 3200. Each spot sample was calibrated by using an external calibration solution. MS/MS analysis was performed using 1 kV collision energy and air as a collision gas. For peptide sequencing, the 10 most intense precursors were selected. Data were acquired automatically after 1500 laser shots with a fixed laser intensity of 4200. Data were analyzed using GPS Explorer software (Applied Biosystem) and MASCOT software (Matrix Science, London, UK). NCBInr and human were selected as the database and taxonomy, respectively. The remaining peptide mixtures were separated by using a nanoflow-HPLC system (Ultimate; Switchos; Famos; LC Packings, Amsterdam, The Netherlands). A sample volume of 10 µl was loaded by the autosampler onto a homemade 2 cm fused silica pre-column (75 µm I.D.; 375 µm O.D.; Resprosil C18-AQ, 3 µm (Ammerbuch-Entringen, DE) at a flow rate of 2 µl/min. Sequential elution of peptides was accomplished using a flow rate of 200 nl/min and a linear gradient from Solution A (2% acetonitrile; 0.1% formic acid) to 50% of Solution B (98% acetonitrile; 0.1% formic acid) in 40 minutes over the pre-column in-line with a homemade 10-15 cm resolving column (75 µm I.D.; 375 µm O.D.; Resprosil C18-AQ, 3 µm (Ammerbuch-Entringen, Germany). Peptides were eluted directly into an ion trap Esquire 3000 plus (Bruker-Daltonik, Germany). 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. Protein identification was performed by searching in the National Center for Biotechnology Information non-redundant database (NCBInr) using the Mascot program (http://www.matrixscience.com). The following parameters were adopted for database searches: complete propionamide formation on cysteines and partial oxidation of methionines, peptide Mass Tolerance ±1.2 Da, Fragment Mass

Tolerance ±0.9 Da, missed cleavages 2. For positive identification, the score of the result of [-10xLog(P)] had to be over the significance threshold level ($P<0.05$).

2.3.2.9 Protein Validation By Western Blot Analysis

For the purpose of this analysis, three biological replicate experiments for the two samples (control and RC307-treated cell line) were analyzed. Protein extracts from the HCT116 cell line treated with RC307 for 24 h (2.5 μM) as described above were diluted 1:1 with Laemmli's sample buffer (62.5 mM Tris-HCl, pH 6.8, 25% glycerol, 2% SDS, 0.01% Bromophenol Blue, 5% ß-mercaptoethanol) boiled for 3 min and separated by SDS/polyacrylamide gel electrophoresis (PAGE) on 12% T acrylamide gels in Tris/glycine/SDS buffer. Protein electroblotting and chemiluminescent signal detection were performed as in [29]. Briefly, 1-D SDS-PAGE gels were transferred to a PVDF membrane and treated with the respective antibodies at the appropriate dilutions. The bound antibody was detected by enhanced chemiluminescent (ECL) detection kit (Amersham Biosciences Europe) and recorded with X-ray X-Omat AR (Kodak, Rochester, NY, USA) films. Membranes were immunoblotted again with a monoclonal anti-β actin antibody (Sigma-Adrich, 1:4.000) for normalization purposes. The intensity of the chemiluminescence response was measured by scanning films and processing the image using Quantity One software Version 4.4 (Bio-Rad).

2.3.2.10 Protein Categorization

Gene ontology (GO) lists were downloaded using the tool FatiGO [30] from Babelomics (http://fatigo.bioinfo.cipf.es/), a complete suite of web tools for the functional analysis of groups of genes in high-throughput experiments. Each protein was classified with respect to its cellular component, biological process, and molecular function using GO annotation. When no GO annotation was available, proteins were annotated manually based on literature searches and closely related homologues.

2.3.3 RESULTS

2.3.3.1 Effects Of RC307 On HCT116 Cells

Sensitivity of HCT116 cells to the HDAC inhibitor RC307 was assessed by growth-inhibition assay following exposure to the drug for 72 h. The compound exhibited a marked capability to inhibit proliferation, the IC_{50} value being in the micromolar range, i.e, 1.14±0.15 µM (n=6) (Fig. 2A). Such an effect was associated with induction of apoptosis as shown by TUNEL assay in cells exposed for 72 h to a concentration of the compound producing an inhibition of cell growth of around 80% (Fig. 2B).

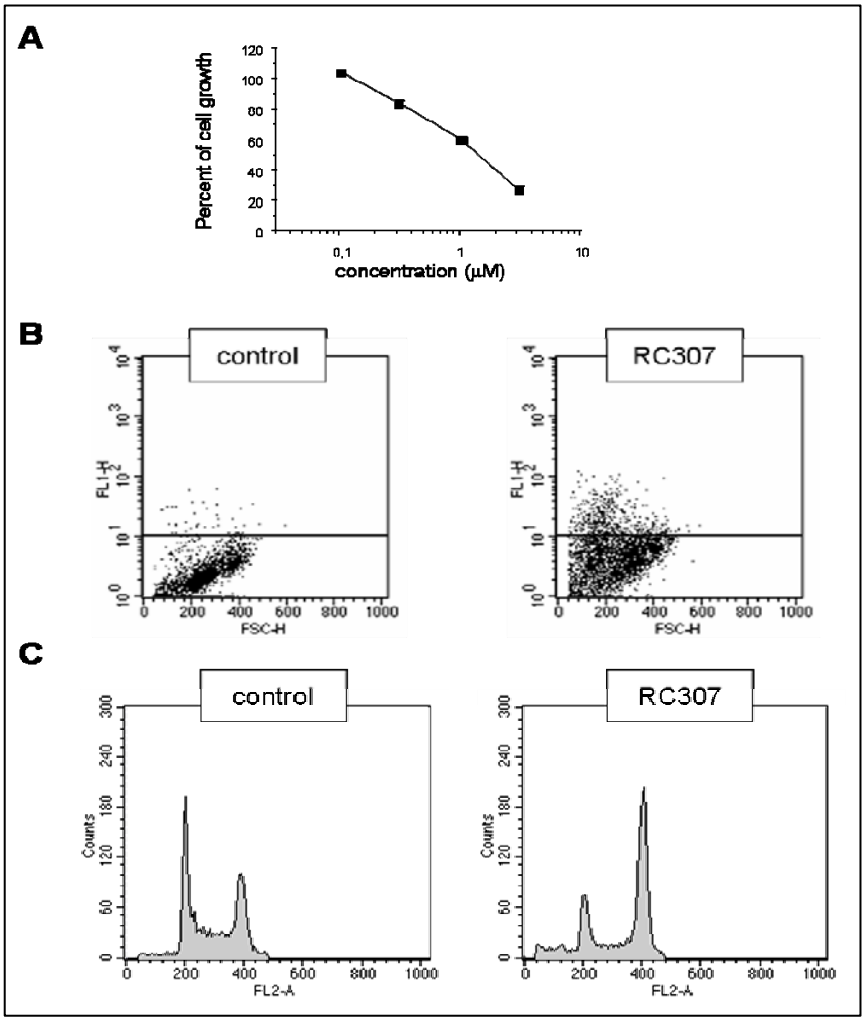

Fig. 2. A. Sensitivity of HCT116 cells to RC307. Sensitivity was assessed by growth inhibition assay with a 72 h exposure and cells were counted at the end of treatment. Values are the mean (±SD) of six independent experiments. B. Analysis of apoptosis in HCT116 cells exposed to RC307. Apoptosis was measured by TUNEL assay after 72 h exposure to a concentration of RC307 corresponding to the IC_{80} (2.5 µM). Representative dot plots showing viable cells ($FL1<10^1$) and apoptotic cells ($FL1>10^1$) are presented. C. Cell cycle distribution in HCT116 cells exposed to RC307. Cell cycle distribution was analysed 24 h after exposure to RC307 (concentration corresponding to IC_{80}) in propidium iodide stained cells. Hystograms from a representative cytofluorimetric analysis are shown.

Indeed, an appreciable amount of cells with DNA fragmentation was documented by labelling of DNA strand breaks. The occurrence of apoptotic cell death was also supported by cytofluorimetric analysis of cell cycle showing the appearance of a sub-G1 peak after 24 h drug exposure. Under such conditions, the antiproliferative treatment produced a marked accumulation in the G2 phase of the cell cycle (Fig. 2C).

2.3.3.2 HCT116 Cell Line 2-DE Protein Pattern Analysis

In order to examine the effect of HDAC inhibition by RC307 on protein expression in a human colorectal carcinoma cell line (HCT116), five replicated maps for each experimental group (controls and RC307-treated), were performed. The separated protein spots were visualized on 2D gels by Sypro Ruby staining, which allows good reproducibility and protein spot quantification for comparison analysis. The 2D gel of total lysates showed a total of 480±21 and 474±22 protein spots, for the control and RC307-treated cells respectively; while the 2D gels of nuclear extracts showed a total of 774±19 and 781±15 protein spots, for the control and RC307-treated cells respectively. By PDQuest analysis of the 2D gel replicates differential protein expression between control and RC307-treated cells was measured. A total of 48 and 46 different spots were found to be modulated by RC307 in total lysates and nuclear extracts, respectively. In particular, 27 spots were found to be up-regulated and 21 down-regulated in the total lysates; while 24 up-regulated and 22 down-regulated in the nuclear extracts. Fig. 3A and Fig. 3B show representative 2D gels of total lysate and nuclear proteome of HCT116 cell line, respectively, together with some of the corresponding differentially expressed spots ($p<0.05$) between control and RC307-treated samples.

Spots selected as regulated from the differential analysis were subjected to MALDI-TOF/TOF and RP-HPLC-ESI-MS/MS analysis for protein identification. The unique differentially expressed proteins identified were 66. In Table 1 the identity of the successfully identified proteins corresponding to up- or down-regulated spots (confidence level of 0.05) is shown, together with the standard spot number (SSP), the MS identification parameters and the indication of their gene ontology (GO) annotation (cellular component, biological process and molecular function). Not all the differentially expressed spots were identified because of their relative low concentrations. Fig. 4

shows as a pie chart the distribution of the identified proteins catalogued according to the biological process in which they are involved.

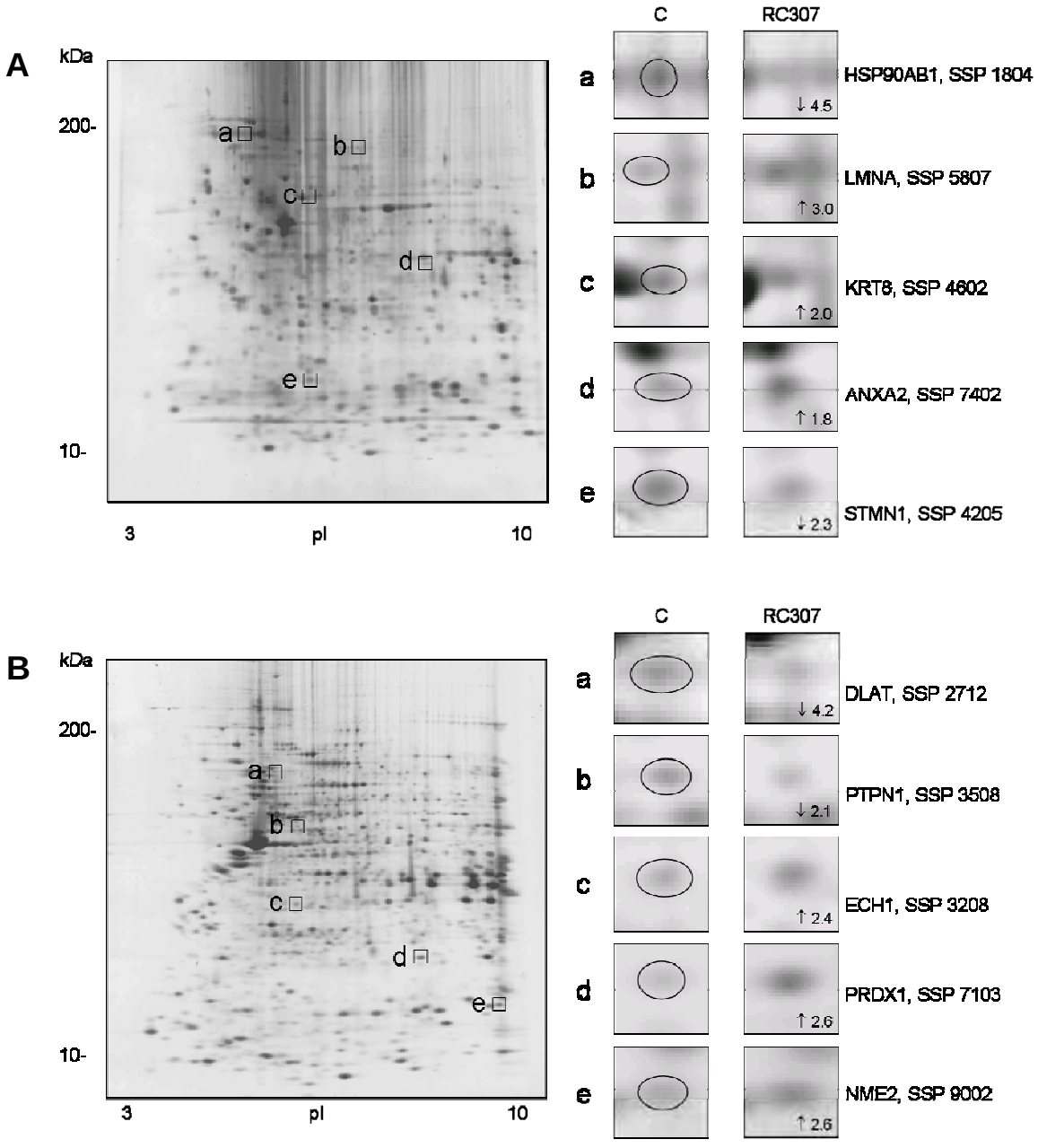

Figure 3. Example of Sypro Ruby stained 2-D gels of (A) total lysates and (B) nuclear proteins extracted from HCT116 cell line, and PDQuest (version 7.3) output showing representative differentially expressed spots ($p<0.05$) between control and RC307-treated samples. For each spot an enlarged region of the respective 2-DE map is shown which contains the referred spot (highlighted by an ellipse) and the corresponding fold of change. Each spot was identified by MS/MS analysis.

Table 2. Identified proteins that are differentially expressed at 24 hours of RC307 treatment through 2-DE.

Protein name	SSP	Extract	HUGO Gene Name	NCBI accession no.	No. of peptides	Mascot Score	Fold of variation in treated cells	Cellular Component GO term	Molecular Function GO term
Cell proliferation, cell cycle and apoptosis									
Adenylate kinase isoenzyme 1	9204	total lysate	AK1	gi|49456961	11	102	up 1,55	cytosol	adenylate kinase activity
Annexin V	2007	nuclear	ANXA5	gi|1421662	4	133	down 2,70	cytosol	calcium binding
Calreticulin	905	total lysate	CALR	gi|48146257	9	108	up 2,01	endoplasmic reticulum	calcium ion binding
Galectin 3	9301	total lysate	LGALS3	gi|126678	4	143	down 1,65	nucleus	immunoglobulin binding
Glutathione S-transferase pi	4202	total lysate	GSTP1	gi|47496669	7	89	up 1,61	cytosol	glutathione transferase activity
Heat shock 70 kDa protein 1 (hsp72)	1706	nuclear	HSPA1	gi|188488	1	71	down 2,78	endoplasmic reticulum	ATP binding, protein binding
Heat shock 70 kDa protein 5 (grp78)	702	nuclear	HSPA5	gi|386758	1	112	up 2,09	endoplasmic reticulum	ATP binding
Heat shock 90kDa protein 1, beta	1804	total lysate	HSP90AB1	gi|20149594	19	756	down 4,52	cytosol	ATP binding
Inosine monophosphate dehydrogenase	4611	nuclear	IMPDH2	gi|44979607	5	244	up 2,03	cytosol	oxidoreductase activity
Parathymosin	201	total lysate	PTMS	gi|46276863	4	231	up 1,54	nucleus	protein binding
Peroxiredoxin 1	7103	nuclear	PRDX1	gi|4505591	13	464	up 2,58	cytosol	antioxidant activity
Rho GDP dissociation inhibitor	1104	nuclear	ARHGDIA	gi|3608	3	123	up 2,65	cytoskeleton	enzyme regulator activity
Septin 2	4505	nuclear	SEPT2	gi|1040689	2	104	down 2,22	cytoskeleton	pyrophosphatase activity
Stathmin 1	4205	total lysate	STMN1	gi|5031851	20	685	down 2,29	cytoskeleton	protein binding
TNF receptor-associated protein 1 (hsp75)	5804	total lysate	TRAP1	gi|1082886	4	180	up 1,96	cytosol	protein binding
Regulation of gene expression									
Alpha-enolase	6901	total lysate	ENO1	gi|39644850	9	124	up 7,18	cytosol, nucleus	transcription factor activity
Alpha-enolase	6902	total lysate	ENO1	gi|39644850	8	94	up 111,05	cytosol, nucleus	transcription factor activity
Alpha-enolase	6903	total lysate	ENO1	gi|39644850	7	85	up 14,76	cytosol, nucleus	transcription factor activity
Alpha-enolase	6904	total lysate	ENO1	gi|39644850	6	67	up 111,90	cytosol, nucleus	transcription factor activity
Alpha-enolase	6603	total lysate	ENO1	gi|39644850	10	106	up 2,17	cytosol, nucleus	transcription factor activity
Cytokine induced protein 29 kDa	4203	nuclear	CIP29	gi|32129199	2	113	up 2,51	nucleus	protein binding
Non-metastatic cells 1, protein	1101	nuclear	NME1	gi|35068	3	96	up 7,34	nucleus	transcription factor activity
Non-metastatic cells 2, protein	9002	nuclear	NME2	gi|1421609	5	288	up 2,63	nucleus	nucleoside diphosphate kinase activity
Nuclease sensitive element-binding protein 1	604	total lysate	YBX1	gi|117938841	13	138	up 1,56	nucleus	nucleic acids binding
Proliferation-associated 2G4, 38kDa	4505	nuclear	PA2G4	gi|4099506	7	336	down 2,22	nucleus	amino peptidase activity
Profilin I	106	nuclear	PFN1	gi|999511	3	86	up 2,59	nucleus	protein binding
TBP-associated factor 15	6804	nuclear	TAF15	gi|119600533	3	96	up 2,38	nucleus	nucleic acid binding, cation binding
Signal transduction									
Annexin III	2207	nuclear	ANXA3	gi|1421662	4	243	up 2,76	cytosol	pyrophosphatase activity
Protein tyrosine phosphatase, non-receptor type 1	3508	nuclear	PTPN1	gi|809208	5	215	down 2,13	endoplasmic reticulum	protein tyrosine phosphatase activity
Chromatin and cytoskeleton organization									
Annexin II	5303	nuclear	ANXA2	gi|4757756	6	388	up 3,04	soluble fraction and plasma membrane	phospholipase inhibitor activity
Annexin II	7402	total lysate	ANXA2	gi|73909156	15	204	up 1,8	soluble fraction and plasma membrane	phospholipase inhibitor activity

Table 3 *(continued)*

Annexin II	7501	total lysate	ANXA2	gi|73909156	12	94	up 1,53	soluble fraction and plasma membrane	phospholipase inhibitor activity
Annexin IV	3202	nuclear	ANXA4	gi|189617	2	125	up 7,03	cytosol	phospholipase inhibitor activity
Beta-actin	1607	nuclear	ACTB	gi|28336	4	226	up 2,65	actin filament	protein binding
Cytokeratin 8	4602	total lysate	KRT8	gi|62913980	14	157	up 2,03	cytoskeleton	protein binding
Cytokeratin 8	3604	total lysate	KRT8	gi|62913980	17	212	up 1,94	cytoskeleton	protein binding
Cytokeratin 8	2606	nuclear	KRT8	gi|181573	23	874	up 16,74	cytoskeleton	protein binding
Cytokeratin 18	1101	nuclear	KRT18	gi|30311	5	258	up 7,34	cytoskeleton	protein binding
Histone H4	7001	total lysate	HIST4H4	gi|51317339	7	88	down 4,91	nucleus	DNA binding
Lamin A/C	5804	total lysate	LMNA	gi|5031875	10	290	up 1,96	lamin filament	protein binding
Lamin A/C	5807	total lysate	LMNA	gi|125962	9	124	up 2,96	lamin filament	protein binding
RNA splicing, processing, and translation									
Cleavage and polyadenylation specific factor 6, 68kDa	2712	nuclear	CPSF6	gi|5901928	1	97	down 4,17	nucleus	RNA binding
hnRNP D	6413	nuclear	HNRPD	gi|181914	2	64	down 2,17	nucleus	RNA and DNA binding
hnRNP D	8414	nuclear	HNRPD	gi|508270	5	153	down 3,70	nucleus	RNA and DNA binding
hnRNP H1	3508	nuclear	HNRPH1	gi|5031753	3	137	down 2,13	nucleus	RNA binding
hnRNP K	3704	nuclear	HNRPK	gi|473911	9	386	down 2,56	nucleus	RNA binding
hnRNP R	2705	nuclear	HNRPR	gi|5031755	2	128	down 2,50	nucleus	RNA binding
Mitochondrial ribosomal protein L12	4203	total lysate	MRPL12	gi|20981709	7	86	down 1,71	mitochondrion	RNA binding
Mitochondrial ribosomal protein S22	4509	total lysate	MRPS22	gi|13633893	14	225	down 1,75	mitochondrion	RNA binding
Nuclear RNA-binding protein, 54-kD	8504	nuclear	NONO	gi|543010	4	163	down 1,79	nucleus	RNA and DNA binding
Nucleolin	702	nuclear	NCL	gi|34534595	5	242	up 2,09	nucleus	RNA and DNA binding
Ribosomal protein L23a	201	total lysate	RPL23A	gi|404015	2	128	up 1,54	ribosome	RNA binding
RNA helicase 2	6808	nuclear	DHX15	gi|2696613	5	150	down 2,78	nucleus	RNA helicase activity
Splicing factor SF3a60	1607	nuclear	SF3A3	gi| 551450	9	533	up 2,65	nucleus	RNA binding
Splicing factor, arginine/serine rich 1	2403	total lysate	SFRS1	gi|5902076	5	254	up 1,73	nucleus	RNA binding
Translation initation factor 3, p40 subunit	4406	total lysate	EIF3S3	gi|3986482	3	91	down 1,39	cytosol	RNA and DNA binding
Translation elongation factor 1, alpha	7508	nuclear	EEF1A1	gi|31092	3	128	up 2,55	cytosol	GTPase activity
U6 snRNA-associated Sm-like protein LSm8	1205	total lysate	LSM8	gi|7706425	2	85	down 1,65	nucleus	RNA binding
Protein folding and degradation									
Antigen NY-CO-10	3704	nuclear	SDCCAG10	gi|3170184	2	107	down 2,56	nucleus	peptidyl-prolyl cis-trans isomerase activity
Heat shock 70 kDa protein 8, hsc70	2705	nuclear	HSPA8	gi|62897129	14	873	down 2,50	nucleus	ATPase activity
Proteasome activator complex subunit 3	4401	total lysate	PSME3	gi|49456449	7	84	down 3,95	cytosol	protein binding
Protein disulfide isomerase-associated 5	7618	nuclear	PDIA5	gi|5803121	4	130	up 3,03	endoplasmic reticulum	protein disulfide isomerase activity
Protein disulfide-isomerase precursor	1703	total lysate	P4HB	gi|15680282	7	61	up 1,76	endoplasmic reticulum	peptidyl-proline 4-dioxygenase activity
Protein disulfide-isomerase precursor	1704	total lysate	P4HB	gi|15680282	8	235	up 1,62	endoplasmic reticulum	peptidyl-proline 4-dioxygenase activity
T-complex protein 1 subunit delta	8707	total lysate	CCT4	gi|38455427	15	202	down 1,64	cytosol	ATP binding
Ubiquitin	3005	nuclear	RPS27A	gi|229532	3	61	up 2,72	cytosol, nucleus	protein binding
Electron and mitochondrial transport									
Vascular H+ ATPase E1 isoform a	6211	nuclear	ATP6V1E1	gi|4502317	6	266	up 2,25	mitochondrion	ATPase activity

Table 1. *(continued)*

Others (glucide, lipid, nucleotide metabolism, etc…)									
Acetyl-Coenzyme A acetyltransferase 1	8412	nuclear	ACAT1	gi|39795296	2	66	up 2,94	mitochondrion	transferase activity
Aldolase A protein	8401	nuclear	ALDOA	gi|28595	1	75	down 2,27	cytosol	aldehyde-lyase activity
C-1-tetrahydrofolate synthase	6801	nuclear	MTHFD1	gi|115206	4	243	down 5,00	mitochondrion	cyclohydrolase activity
Dihydrolipoamide S-acetyltransferase	2712	nuclear	DLAT	gi|35360	4	169	down 4,17	mitochondrion	transferase activity
Enoyl-CoA hydratase	3208	nuclear	ECH1	gi|70995211	8	612	up 2,38	mitochondrion	isomerase activity
Phosphoserine aminotransferase	7503	total lysate	PSAT1	gi|20141815	4	62	down 2,03	cytosol	transferase activity
Phosphoserine aminotransferase	7504	total lysate	PSAT1	gi|16741698	8	247	down 1,62	cytosol	transferase activity
Miscellaneous									
Autophagy-related protein 16-1	2505	total lysate	ATG16L1	gi|62510482	7	61	down 2,87	autophagic vacuole	protein binding
DAZ associated protein 1	8503	nuclear	DAZAP1	gi|8671754	3	128	up 2,02	nucleus	RNA binding
Androgen-regulated protein 2	2205	total lysate	HN1	gi|7705877	2	53	up 1,80	nucleus	RNA binding
Myosin regulatory light chain 2	107	nuclear	MYL9	gi|20141521	2	151	up 2,22	cytoskeleton	calcium binding

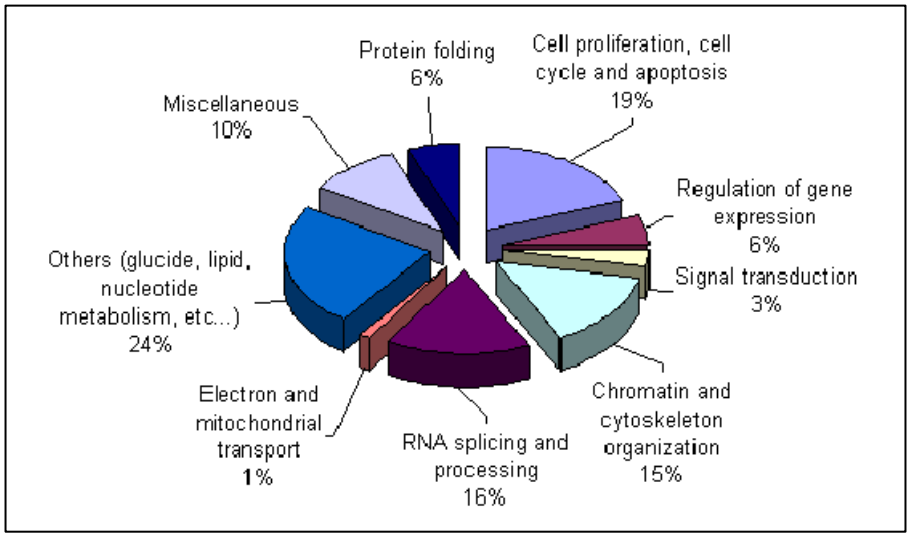

Figure 4. Distribution of the identified proteins according to the biological process in which they are involved. Assignments were made on the basis of information provided by Gene ontology (GO) lists downloaded using the tool FatiGO 13 from Babelomics (http://fatigo.bioinfo.cipf.es/).

To validate the findings obtained by 2-DE, the regulation in level of expression of four candidate proteins (cytokeratin 8, cytokeratin 18, alpha enolase and stathmin) were further investigated by immunoblot analysis. The western blot results are shown in Fig. 5, including the relative protein expression in both samples normalized to β-actin signal intensity as an internal loading control. Trends of changes in the same direction as those detected in the 2D gel analyses were detected for all the four proteins. The quantitative difference between the results obtained by 2D electrophoresis and by Western blot suggested that most changes detected by the former technique

specifically involve post-translationally modified forms, which can be only separated in 2D maps (Fig. 3A and 3B).

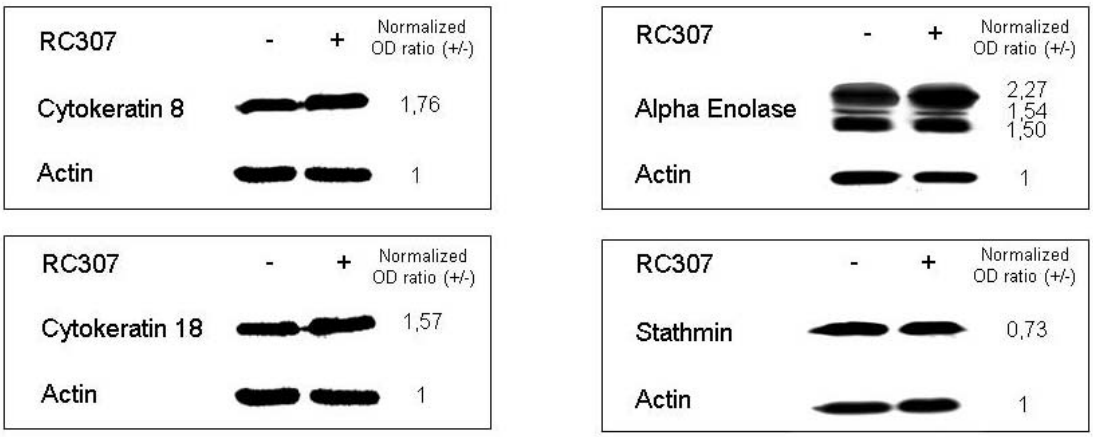

Figure 5. Western blot analyses for five proteins showing regulation in RC307 treated HCT116 cell line by proteomic analysis. Film images with the relative protein expression normalized to actin signal intensity as an internal control. Western blot images were captured by GS710 densitometer (BioRad) and analyzed by QuantityOne software to calculate the band intensities (OD).

2.3.4 DISCUSSION

In this study, it was found that RC307, a histone deacetylase inhibitor, determined cell growth inhibition, G2 accumulation and a moderate induction of apoptosis of HCT116 colon cancer cell line. To analyze the molecular mechanisms involved in cell growth inhibition and the potential target genes for aberrant acetylation, a proteomic analysis of HCT116 after treatment with RC307 was performed. The differentially expressed proteins identified by 2-D proteome analysis are related to various cellular programs involving e.g., proliferation, cell cycle and apoptosis regulation, gene expression, as well as chromatin and cytoskeleton organization (Table 1). Interestingly, some of these identified proteins (such as Glutathione S-transferase pi, Peroxiredoxin 1, TNF receptor-associated protein 1, Profilin I, Annexin II, Cytokeratin 8 and 18, Lamin A/C, and U6 snRNA-associated Sm-like protein LSm8) correspond to those which were reported as modulated by trichostatin A (showing the same trend of variation) in previous studies [23-25].

The cellular effect of RC307 could be explained by modulation of expression of proteins that have been implicated in cell proliferation, cell cycle and apoptosis. In the present study, Galectin 3 (LGALS3) was down-regulated RC307. LGALS3 is a beta-galactoside-

binding protein whose expression has been correlated with progression and metastasis in colon cancer and has been reported to represent a potent prognostic marker in colorectal cancer [31, 32]. Furthermore, the interaction between LGALS3 and Bcl-2 suggests that LGASL3 is involved in the inhibition of apoptosis [33, 34]. It is possible that RC307-mediated LGALS3 reduction could help to induce the apoptotic response of HCT116 cells. RC307 also modulates the Glutathione S-transferase pi (GSTP1), which belongs to a family of isoenzymes known to inactivate damaging electrophilic compounds by catalyzing their conjugation to reduced glutathione. In most experimental systems, over-expression of GSTP1 is associated with increased resistance to anticancer agents [35]. Although there are contrasting reports regarding the association of GSTP1 up-regulation with the apoptotic response induced in colon cancer cells by the HDAC inhibitor butyrate [36, 37], it has previously been shown that the apoptotic response induced in pancreatic cancer cell lines by TSA is related to GSTP1 over-expression [24, 25]. These findings are intriguing, since the lack of GSTP1 expression is associated with methylation in the gene promoter [38], and methylation-induced gene silencing has been associated with histone deacetylation. Findings in this work showing up-regulation of GSTP1 by RC307 are in keeping with such reports, as inhibition of HDAC could relieve gene silencing. Another protein modulated by RC307 was Peroxiredoxin 1 (PRDX1), a ubiquitously expressed member of a family of antioxidant proteins induced by reactive oxygen species. PRDX1 has been reported to inhibit the c-myc oncogene and to act as a tumour suppressor [39, 40]. Moreover, other HDAC inhibitors (i.e., FK288 and TSA) have been shown to up-regulate PRDX1 in esophageal and pancreatic cancer cell lines in which apoptosis has been documented [41, 23-25]. Consistent with this, here it has been shown that also RC307 enhances the level of the tumour suppressor PRDX1. Another protein that was found to be up-regulated by RC307 was the Heat shock 90 kDa protein 1, beta (HSP90AB1), an ATP-dependent chaperone that plays a central role in regulating the stabilization, activation, and degradation of a range of proteins including the products of oncogenes. HSP90AB1 has become an attractive target for novel cancer therapeutic agents, since its inhibition disrupt multiple cancer-causing pathways simultaneously [42, 43]. Recently, it has been reported HCT116 cell growth inhibition after treatment with different hsp90 inhibitors [44]. Since here it was found that RC307 reduces the level of HSP90AB1 expression it

is conceivable that the event is a consequence of HSP90ABI acetylation. Another protein regulated by RC307 was Stathmin (STMN1), a microtubule destabilizing protein previously described as being negatively regulated by p53 [45] and highly expressed in several tumours. Recently, it has been demonstrated that down-regulation of STMN1 by siRNA in osteosarcoma cell lines induces G(2)/M cell cycle arrest and apoptotic cell death [46]. Accordingly, the reduction of STMN1 level increases the responsiveness of tumour cells to treatment with chemotherapeutic agents [47]. Consistent with this, RC307 down-regulates STMN1, and this finding has been validated by immunoblot analysis. Collectively, all these findings suggest that cell response of HCT116 cells to RC307 involves proteins that play a key role in colon cancer cell proliferation, cell cycle arrest and apoptosis.

The cellular effect of RC307 could also be explained by the modulation of proteins involved in gene expression regulation. Alpha-enolase (ENO1) was another protein regulated by RC307. ENO1 is a bifunctional gene encoding both a glycolytic enzyme (ENO1, 48kDa), and a transcription factor (MBP-1, 37kDa) which binds and represses c-myc gene playing an important role in cancer cell growth inhibition [48]. Surprisingly, it has been demonstrated that also the longer form (ENO1) alone has equal or stronger effect on the apoptosis induction, as MBP-1 does [49]. Consistent with this, RC307 was found to up-regulate ENO1, a finding that was validated by immunoblot analysis. Here it was reported that RC307 modulates also cytokine-inducible 29-kDa protein (CIP29). CIP29 is a cytokine regulated nuclear protein that binds both double- and single-stranded DNA. This nuclear protein is involved in normal and cancer cell proliferation [50]. Interestingly, it has been reported that CIP29 has a growth inhibitory effect associated with induction of apoptosis [51]. Accordingly, RC307 was found to up-regulate CIP29 in colon cancer cell lines. It is possible that RC307-mediated ENO1 and CIP29 up-regulation could enhance the apoptotic response of HCT116 cells.

The mechanism of action of RC307 also involves the regulation of proteins related to chromatin and cytoskeleton organization. In particular, among the cytoskeleton related proteins, cytokeratins isoforms 8 (KRT8) and 18 (KRT18) were up-regulated by RC307. KRT8 and KRT18 are the major components of intermediary filaments of simple or single layer epithelia, such as those of the intestine. While both isoforms are essential for maintaining structural integrity, there is accumulating evidence indicating that they

also exert non-mechanical functions. Indeed, the expression of KRT8 and KRT18 has been related to poor clinical prognosis [52]. Moreover, KRT8 and KRT18 up-regulation has been associated with the sensitization to apoptosis, induced by cisplatin [53], roscovitine [54], and Fas and TNF [55]. It has also been reported that KRT8 and KRT18 form intermediate filaments that are required by hsp90β (HSP90AB1) for its chaperone activity [56]. Surprisingly, KRT8 and KRT18 have been shown to be up-regulated in HCT116 cell line after inhibition of the HSP90AB1 [57]. Based on such reports, it is conceivable that increases in KRT8 and KRT18 level of expression here reported, are related to the RC307-mediated HSP90AB1 down-regulation, and may contribute to HCT116 cellular growth inhibition induced by the HDAC inhibitor RC307.

2.3.5 CONCLUSION

In conclusion, the results obtained indicate a modulation of proteins belonging to different pathways (i.e., proliferation, cell cycle, apoptosis regulation, gene expression, as well as chromatin and cytoskeleton organization) by a novel HDAC inhibitor. Among the most interesting RC307-regulated proteins, Galectin 3, Peroxiredoxin 1, Heat shock 90kDa protein 1 beta, Alpha-enolase and Cytokeratins 8 and 18 were identified. These findings support the interest of proteomic approaches in defining the mechanism of action of novel anticancer drugs.

2.4 APPLICATION OF PARTIAL LEAST SQUARES DISCRIMINANT ANALYSIS AND VARIABLE SELECTION PROCEDURES: A 2D-PAGE PROTEOMIC STUDY

2.4.1 INTRODUCTION

The large amount of information contained in groups of 2-DE maps, belonging to the same specimen or to specimens to be compared (e.g. control and pathological subjects), is usually extracted by means of dedicated software packages (PDQuest, Melanie, Progenesis etc.). The classical differential analysis thus carried out is based on the direct analysis of images of maps, previously scanned via a densitometer. The multi-step procedure consists briefly in: 1) background correction of the maps to be analysed; 2) alignment of the maps; 3) identification of the spots present independently on each map (fitting with Gaussian distributions); 4) matching of the spots present on all the maps belonging to the same group (e.g. control or treated samples); 5) achievement of the “master gel” for each class containing the spots present contemporarily on all the maps belonging to the same group.

The final differential analysis can be carried out by the same software, which will ultimately perform a statistical test (e.g. Student’s t-test) and list the spots responsible for the differences between the master gels of each group of samples. In alternative, a spot volume dataset can be obtained where each sample (map) of each group is described in terms of the volumes (sum of optical densities) of all the spots identified, namely those contained on all the maps of that group contemporarily and those contained in a restricted subset of maps. Such datasets, where a certain number of samples is described by a large number of variables (spots), can be effectively treated by means of multivariate statistical tools, such as Principal Component Analysis (PCA), classification methods [Linear Discriminant Analysis (LDA), Soft-Independent Model of Class Analogy (SIMCA)] and cluster analysis techniques. Multivariate tools are widely applied to proteomic datasets [58-66]. In fact, while univariate methods (such as Student’s t-test) are used to detect significant changes in the expression of individual proteins, the multivariate methods look for patterns in expression changes and utilise all the data simultaneously. Some recent applications of PCA, SIMCA and other

multivariate tools were presented by Prof. Marengo (Department of Environmental and Life Sciences, Eastern Piemonte University, Alessandria) [67-70]. Other papers report the use of Partial Least Squares – Discriminant Analysis (PLS-DA) as a classification tool for identifying the differences between groups of samples [71-78]. In the present work (performed in collaboration with Prof. Marengo group), PLS-DA was applied coupled to variable selection procedures (backward elimination – BE) for identifying the actually significant spots responsible for the differences existing between the groups of samples in two different datasets: nuclei from HCT116 (6 control and 5 RC307-treated samples) and total cell lysates from HCT116 (5 control and 5 RC307-treated samples), the same of the study described in 2.3. The results obtained from the statistical analyses were then compared to those provided by a classical PDQuest differential analysis based on individual Student t-tests (α=0.05). The PLS-DA analysis applied in this study permitted to validate and improve the PDQuest analysis.

2.4.2 THEORY

2.4.2.1 Partial Least Squares-Discriminant Analysis (PLS-DA)

Partial Least Squares (PLS) [79-83] is a multivariate regression method establishing a relationship between one or more dependent variables (Y) and a group of descriptors (X). *X*- and *Y*-variables are modelled simultaneously, to find the latent variables (LVs) in *X* that will predict the latent variables in *Y*. These LVs (also called PLS-components) are similar to the principal components calculated from Principal Component Analysis [79, 80]: they are computed hierarchically, i.e. the first component accounts for the largest amount of information, followed by the other components, accounting for the maximum residual variance. If several responses are presents (more than one Y variable), they are modelled together by the so-called PLS2 algorithm. Since LVs are computed hierarchically, the last LVs are mostly responsible for random variations and experimental error. The optimal number of LVs, i.e. those modelling information in X useful to predict the response Y but avoiding over-fitting, is determined on the basis of the residual variance in prediction. Here, leave-one-out cross-validation is applied to

evaluate the predictive ability and select the optimal number of latent variables on X and Y contained in the final model.

In the case where a large number of descriptors (X variables) are present or a large experimental error is expected, it can be quite difficult to obtain a final model with a suitable predictive ability. In these cases, techniques for variable selection are usually exploited. Here, two subsequent strategies were applied: a first simplification of the model by eliminating groups of non-significant X variables, reducing by about 50% the number of variables present, according to the minimum error in cross-validation; a second phase where variables were eliminated one-at-a-time to provide the final model with the overall minimum error in cross-validation (backward elimination procedure – BE).

PLS was originally set up to model continuous responses but it can be applied even for classification purposes by establishing an appropriate Y related to the association of each sample to a class. Since two classes are present in this case (control and treated samples, in two different datasets), a binary Y variable is added to each dataset, coded so that -1 is attributed to control samples and +1 to treated samples. The regression is then carried out between X-block variables (spot volumes) and the Y just established. This application for classification purposes is called PLS – Discriminant Analysis (PLS-DA).

2.4.2.2 Models Evaluation

The coefficient of multiple determination, R^2, for PLS was calculated as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \qquad (1)$$

where the two sums run on the samples used for calibration (R^2), or for validation (R^2_{CV});

$\hat{y}_i$: predicted value of the response for the *i*-th experiment;

$\bar{y}$: average response of the samples used for calibration (R^2), or for validation (R^2_{CV}).

The root mean square error (RMSE) is calculated as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n}} \qquad (2)$$

where the sum runs on the samples used for calibration (RMSEC), for validation (RMSECV) or for genuine prediction (RMSEP);

$\hat{y}_i$: predicted value of the response for the *i*-th experiment.

In this case, the best model complexity in regard to both the number of X variables present and the number of latent variables in the model is selected by the minimum value of RMSECV (leave-one-out procedure). RMSEP was not calculated since no separation of the samples in training, evaluation and production sets was possible, due to the small number of samples present in both datasets.

2.4.3 MATERIALS AND METHODS

2.4.3.1 Datasets

The analysis was carried out independently for the two cases under investigation, i.e. nuclei and lysates datasets. Nuclei dataset consists of 11 samples (6 control – C – and 5 treated samples – T) described by 779 spot volumes. The dataset deriving from lysates consists instead of 10 samples (5 control – C - and 5 treated samples - RC) described by 528 spots. For both datasets, a binary Y variable was added, representing the belonging of each sample to a class: -1 was attributed to control samples and 1 to treated samples. For both cases, each spot is univocally identified by its SSP number (Standard SPot number).

2.4.3.2 Software

Differential analysis was performed by PDQuest (v.7.3 Bio-Rad), as described in 2.3.2.6. Multivariate analysis and graphical representations were performed by Statistica v.7.1 (Statsoft Inc., USA), The Unscrambler v.9.5 (CAMO, Norway), Parvus 2006 (kindly provided by Prof. Michele Forina, University of Genova, Italy) and Microsoft Excel (Microsoft Corporation, USA).

2.4.4 RESULTS AND DISCUSSION

The human colon cancer HCT116 cell line treated with the new HDAC inhibitor RC307 was chosen as a model for comparing the classical PDQuest image analysis to the PLS-DA coupled to variable selection procedures. The spot volume datasets provided by PDQuest were analysed by BE-PLS-DA to identify the up- and down-regulated spots in both datasets. The results obtained were compared to classical PDQuest differential analysis.

As already pointed out in the theory section, variable selection strategies can be applied to regression and classification problems to identify the best subset of variables providing a model characterised by the best predictive performance.

In this case, the backward elimination algorithm was applied to both datasets, following a two steps procedure:

- groups of variables were first eliminated; in each step no more than 6% of the predictors were eliminated. A first reduction of the dataset of more than 50% was finally obtained.
- predictors were then eliminated one at a time. The model containing the number of predictors able to provide the minimum error in cross-validation (RMSECV) was selected.

2.4.4.1 Nuclei Dataset

- *BE-PLS-DA*

The dataset of dimensions 11 x 779 (11 being the samples and 779 being the variables: 778 spot volumes and the Y variable) was autoscaled before PLS-DA. The autoscaling procedure is particularly important in proteomic datasets where large scale effects are present among the variables, allowing all variables to account for the same amount of variance. In this way also differences between groups due to variations in small spots can be effectively detected.

The first screening by the backward elimination procedure allowed to reduce the dataset to 207 predictors. Then predictors were eliminated one at a time: the best

predictive ability was obtained with a final model containing 61 variables and the first latent variable.

The percentage of explained and cumulative explained variance of the Y and X variables are reported in table 2(a) for the first 9 LVs calculated. The first LV was retained as significant since it explains about 80% of the variance contained in the X block and more than 99% on Y.

Table 4. Variance explained by the LVs calculated for x and y variables for nuclei (a) and lysates dataset (b).

	(a) NUCLEI DATASET					**(b) LYSATES DATASET**			
	BE-PLS-DA		**PDQuest PLS-DA**			**BE-PLS-DA**		**PDQuest PLS-DA**	
	X	Y	X	Y		X	Y	X	Y
	% Expl Var	% Expl Var	% Expl Var	% Expl Var		% Expl Var	% Expl Var	% Expl Var	% Expl Var
LV 1	80.18	99.99	68.12	99.46	LV 1	50.93	99.88	60.76	97.56
LV 2	5.89	0.01	8.71	0.38	LV 2	7.13	0.10	6.24	2.11
LV 3	3.04	0.002	7.58	0.12	LV 3	6.21	0.01	5.30	0.26
LV 4	1.57	0.0001	3.60	0.034	LV 4	6.83	0.001	4.50	0.06
LV 5	2.24	0.00002	1.81	0.011	LV 5	5.26	0	4.90	0.01
LV 6	1.33	0	2.63	0.002	LV 6	5.60	0	5.68	0.001
LV 7	1.16	0	2.27	0	LV 7	5.49	0	4.55	0
LV 8	1.67	0	1.13	0	LV 8	6.38	0	4.76	0
LV 9	1.60	0	2.14	0					

The score and loading plots of the first two LVs are represented in Fig 6(a). The score plot represents the samples well separated in control (at negative values on LV_1, represented as circles) and treated samples (at positive values on the same LV, represented as squares). The analysis of the loading plot allows to identify at positive loadings on LV_1 the spots more intense in treated samples (up-regulated after the treatment) and, at negative loadings on LV_1, the spots more intense in control samples (down-regulated after the treatment).

The regression coefficients calculated on the basis of the first significant LV are represented in Fig. 7(a) on a “virtual” map on a colour scale. Each spot is represented as a circle centred in the x-y position (identified by PDQuest) and identified by its SSP number. Spots are represented on a colour scale according to their increasing positive (from light to dark red) or negative (from light to dark blue) coefficient. Spots

characterised by large negative coefficients are thus more intense in control samples (down-regulated after the treatment), while spots characterised by large positive coefficients are more intense in treated samples (up-regulated after the treatment).

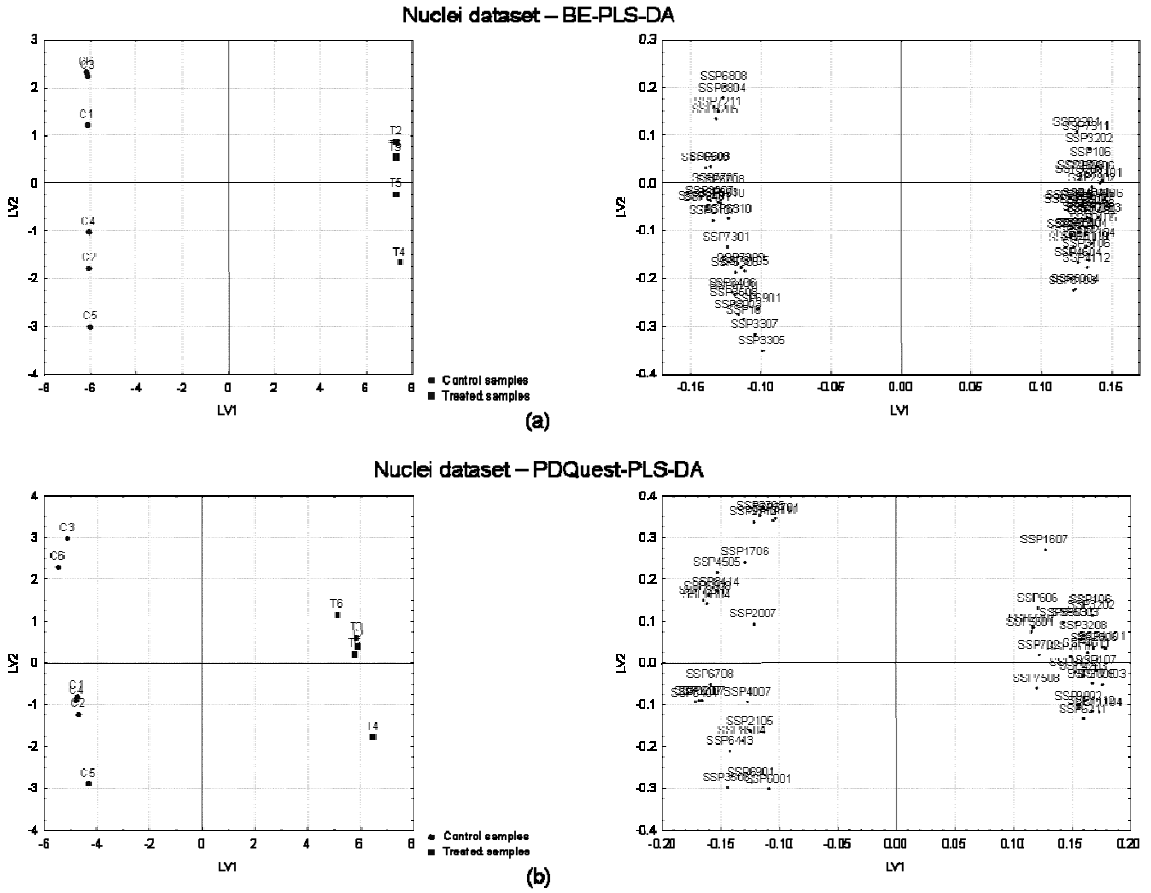

Fig. 6. Nuclei dataset: score plot and loading plot (a) of the first two latent variables for the final PLS-DA model with variable selection; score plot and loading plot (b) of the first two latent variables for the PLS-DA model containing the spots identified by PDQuest. Score plots: control samples are represented as circles, treated samples as squares; Loading plots: X variables are represented as circles, the Y variable as a square.

Table 3(a) reports the performances of the model built with 61 predictors and the first LV as the significant one, both in fitting and prediction.

The model shows very good performances for what regards both its ability to fit data and its capability in prediction, according to the values of R^2 calculated. For what regards the error committed by the model (RMSE), it is extremely low in fitting (about 0.0107) and very good in prediction (about 0.0192). These results are also confirmed by the representation of the calculated vs. measured responses for both calibration and validation samples (Fig. 8(a)), identifying extremely low variations along the y-axis.

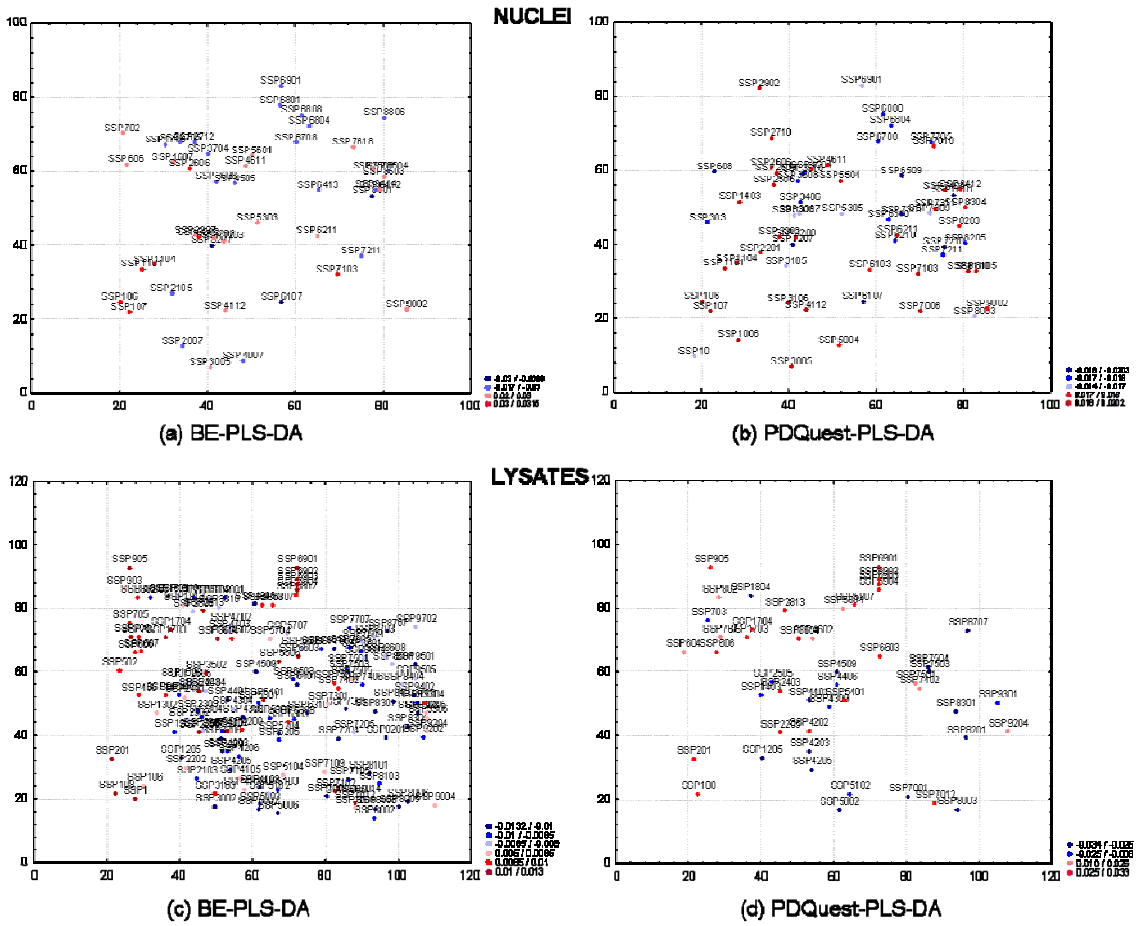

Fig. 7. Nuclei dataset: regression coefficients represented for the final BE-PLS-DA (a) and PDQuest-PLS-DA (b) models. Lysates dataset: regression coefficients represented for the final BE-PLS-DA (c) and PDQuest-PLS-DA (d) models. Each spot is represented on a virtual map as a circle centred in the x-y position identified by PDQuest. The coefficients are represented on a colour scale: from light to dark blue correspond to increasing negative coefficients; from light to dark red correspond to increasing positive coefficients.

Table 5. Performance of BE-PLS-DA and PDQuest-PLS-DA models in fitting and prediction for Nuclei (a) and Lysates (b) datasets (1 LV in all models).

	(a) NUCLEI DATASET			
	BE-PLS-DA		PLS-DA on PDQuest	
	Calibration	Validation	Calibration	Validation
R^2	0.9999	0.9999	0.9946	0.9888
RMSE	0.01070	0.01925	0.08114	0.1091
	(b) LYSATES DATASET			
	BE-PLS-DA		PLS-DA on PDQuest	
	Calibration	Validation	Calibration	Validation
R^2	0.9988	0.9980	0.9756	0.9537
RMSE	0.03846	0.04549	0.1745	0.2253

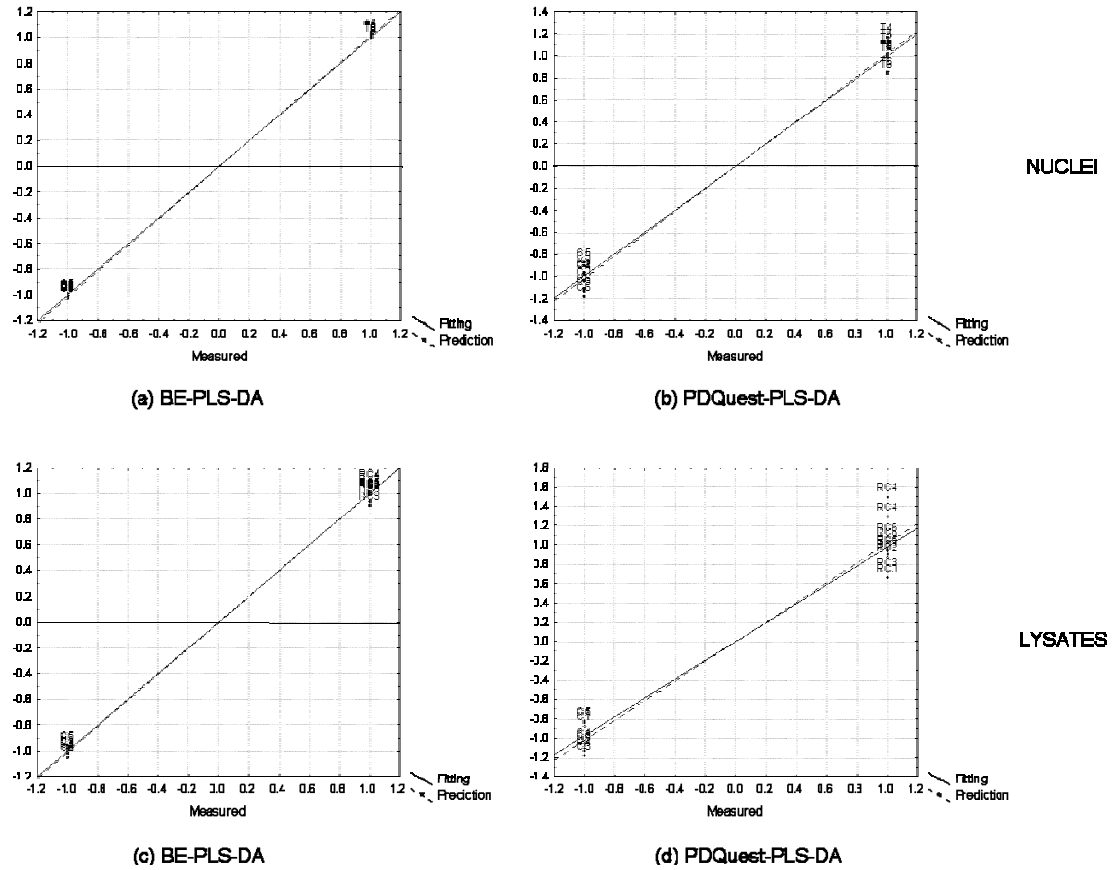

Fig. 8. Calculated Y vs. measured Y in fitting and prediction: for NUCLEI dataset for the final PLS-DA model with variable selection (1 LV in the model) (a) and for the PLS-DA model based on the spots identified by PDQuest (b); for LYSATES dataset, for the final PLS-DA model with variable selection (1 LV in the model) (c) and for the PLS-DA model based on the spots identified by PDQuest (d). The calibration samples are represented as circles, the validation samples as squares.

- *PLS-DA on PDQuest spots – Comparison to BE-PLS-DA model*

For comparison, a PLS-DA model was built considering the 46 spots identified as the relevant ones by PDQuest as the X-block and the Y variable already coded for the previous case as the dependent variable. The data were autoscaled before performing PLS-DA. Table 2(a) reports the percentage of explained variance for the first 9 LVs calculated, compared to that explained by the first 9 LVs calculated by BE-PLS-DA. The first LV appears to be the significant one since it explains more than 99% of the information contained in the Y variable but only 68% in the X block. Fig. 6(b) represents score and loading plots for the first two LVs. As for the previous case, the two groups of samples are clearly separated along the first LV: control samples (circles) at negative values and treated samples (squares) at positive scores along LV_1. The corresponding loading plot identifies 22 spots at negative values along LV_1, more intense in control

samples (down-regulated after the treatment) and 24 spots at positive values along the same LV, more intense in treated samples (up-regulated after the treatment).

The coefficients calculated with the first LV in the final model are represented in Fig. 7(b) on a “virtual” map on a colour scale, as for the previous case. Looking at the results reported in table 3(a), the performances of this model appear good, according to the values of R^2 calculated both in fitting and in validation; the same conclusions can be drawn even if the error committed by the model is examined. However, if the performances of this last model are compared to those of the BE-PLS-DA model, it is clear that some useful information is not taken into account by PDQuest differential analysis. The lower values of R^2 in fitting (and the consequent larger value of RMSE in calibration) show that the PLS-DA model on PDQuest spots is characterised by a smaller fitting ability. The performances of this last model appear even worse if the results in validation are considered, obtaining a final error about 5 times larger than that calculated by BE-PLS-DA model. The same conclusions can be drawn from Fig. 8(b), representing the calculated vs. measured responses both in fitting and prediction: larger deviations can be identified along the y-axis if compared to the BE-PLS-DA model (Fig. 8(a)).

The comparison between the significant spots identified by PDQuest differential analysis and those identified by BE-PLS-DA, shows that the PLS-DA procedure with variable selection identifies about 52.17% of the spots identified also by PDQuest (24 of 46); some of them are identified exclusively by PDQuest (47.83% - 22 of 46) and 37 spots are identified only by means of BE-PLS-DA. BE-PLS-DA proves thus to be more effective than standard differential analysis in the identification of the relevant spots, identifying a larger number of significant variables and providing a smaller error in prediction. The spots identified by the two procedures, divided into up- and down-regulated after the treatment, together with the values of the coefficients calculated by the BE-PLS-DA model, are reported in table 4, according to their decreasing positive or negative coefficient.

Table 6. Nuclei dataset: BE-PLS-DA coefficients of the model containing the first LV.

Down-Regulated										Up-Regulated									
SSP	**Coeff.**	**Both**	**BE**	**PDQuest**	**SSP**	**Coeff.**	**Both**	**BE**	**PDQuest**	**SSP**	**Coeff.**	**Both**	**BE**	**PDQuest**	**SSP**	**Coeff.**	**Both**	**BE**	**PDQuest**
SSP7210	-0.0203		X		SSP10	-0.0161		X		SSP1101	0.0201	X			SSP2506	0.0186		X	
SSP6509	-0.0198		X		SSP3105	-0.0159		X		SSP7006	0.0201		X		SSP2608	0.0183		X	
SSP8401	-0.0197	X			SSP3307	-0.0150		X		SSP7103	0.0200	X			SSP6211	0.0182	X		
SSP3207	-0.0194	X			SSP6901	-0.0146	X			SSP107	0.0196	X			SSP3208	0.0182	X		
SSP608	-0.0194		X		SSP3305	-0.0142		X		SSP8105	0.0195		X		SSP4604	0.0180		X	
SSP6107	-0.0192	X			SSP1706				X	SSP1403	0.0195		X		SSP2201	0.0180		X	
SSP7705	-0.0188		X		SSP2007				X	SSP2902	0.0195		X		SSP5004	0.0176		X	
SSP8205	-0.0188		X		SSP2105				X	SSP1104	0.0195	X			SSP9002	0.0176	X		
SSP303	-0.0186		X		SSP2705				X	SSP2606	0.0193	X			SSP8203	0.0174		X	
SSP7211	-0.0186	X			SSP2712				X	SSP106	0.0193	X			SSP6103	0.0174		X	
SSP6708	-0.0183	X			SSP3704				X	SSP1006	0.0193		X		SSP8304	0.0174		X	
SSP3510	-0.0183		X		SSP4007				X	SSP7408	0.0192		X		SSP7618	0.0172	X		
SSP6804	-0.0181	X			SSP4505				X	SSP3202	0.0192	X			SSP8412	0.0170	X		
SSP6808	-0.0180	X			SSP6413				X	SSP3005	0.0191	X			SSP1607				X
SSP7301	-0.0177		X		SSP6801				X	SSP4611	0.0189	X			SSP2207				X
SSP6310	-0.0177		X		SSP8414				X	SSP2710	0.0189		X		SSP4203				X
SSP3406	-0.0173		X		SSP8504				X	SSP7311	0.0189		X		SSP5303				X
SSP6210	-0.0171		X		SSP8806				X	SSP4112	0.0188	X			SSP5601				X
SSP3508	-0.0171	X								SSP3106	0.0188		X		SSP606				X
SSP5305	-0.0170		X							SSP8110	0.0187		X		SSP702				X
SSP8003	-0.0166		X							SSP5504	0.0186		X		SSP7508				X
SSP7309	-0.0163		X							SSP2506	0.0186		X		SSP8503				X

2.4.4.2 Lysates Dataset

- *BE-PLS-DA*

The dataset deriving from lysate analysis, of dimensions 10 x 529 (10 being the samples and 529 the variables: 528 spot volumes and one Y variable), was autoscaled before the application of PLS-DA. The first step of the backward elimination procedure provided a dataset of 227 predictors. Then, the variables were eliminated one at a time for identifying the model with the best predictive ability. The final model contains 146 predictors and only the first latent variable. Table 2(b) reports the percentage of explained and cumulative explained variance for both Y and X variables for the first 8 LVs calculated. The first LV was considered the relevant one, since it accounts for more than 99% of the variance contained on the Y variable, even if it is related to only 50% of variance on the X block.

Fig. 9(a) represents the score plot and loading plot of the first two LV calculated. From the score plot, it is clear that the first LV is effective in the separation of the two groups

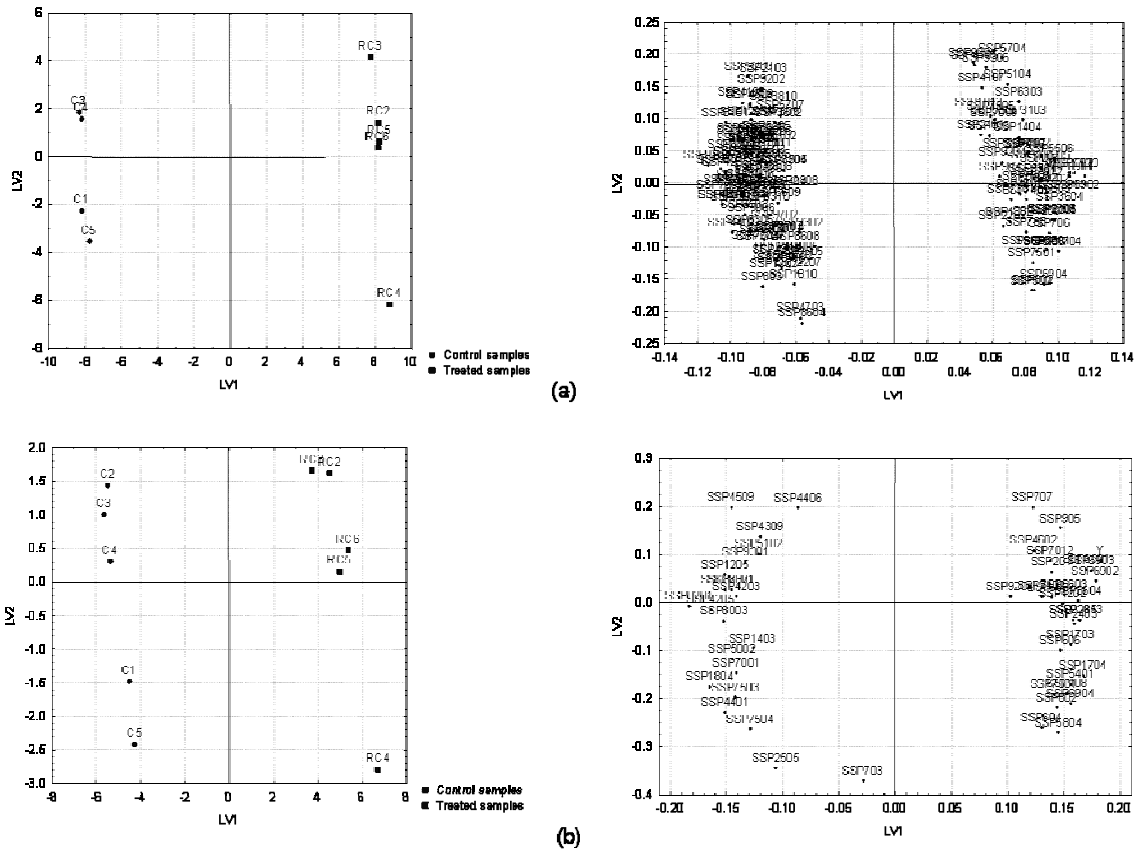

Fig. 9. Lysates dataset: score plot and loading plot (a) of the first two latent variables for the final PLS-DA model with variable selection; score plot and loading plot (b) of the first two latent variables for the PLS-DA model containing the spots identified by PDQuest. Score plots: control samples are represented as circles, treated samples as squares; Loading plots: the X variables are represented as circles, the Y variable as a square.

of samples: control samples (circles) at negative values and treated samples (squares) at positive ones.

The corresponding loading plot allows the identification of two groups of spots: the spots present at negative values along LV_1 are more intense in control samples (down-regulated after the treatment), while those at positive loadings on the same LV are more intense in treated samples (up-regulated after the treatment).

Fig. 7(c) represents the regression coefficients calculated on the basis of the first LV in the final model, represented on a "virtual" map on a colour scale. Again, spots are represented as circles located on the virtual map by their x-y coordinates identified by PDQuest analysis: the increasing red tones correspond to increasing positive coefficients (spots up-regulated after the treatment), while the increasing blue tones correspond to increasing negative coefficients (spots down-regulated after the treatment).

According to the results reported in table 3(b), the performances of the model are very good both in fitting and in prediction, allowing to obtain similar RMSE values in calibration and prediction (smaller than 0.05). The results are confirmed also by Fig. 8(c), representing the measured vs. calculated response in fitting and prediction: the graph shows in fact very small variations along the y-axis.

- *PLS-DA on PDQuest spots – Comparison to BE-PLS-DA model*

As for nuclei dataset, the results obtained by BE-PLS-DA were compared to those obtained by the classical differential analysis by PDQuest. PLS-DA was then applied to the dataset described by the 48 spots identified as relevant by PDQuest analysis. Table 2(b) reports the percentage of explained and cumulative explained variance on Y and X variables, calculated for the first 8 LVs. The first LV appears to be the significant one, explaining about 97% on the Y variable and more than 60% on the X block. The first LV was then selected as relevant and was included in the final model.

Fig. 9(b) represents the score plot and loading plot of the first two LVs; the score plot shows the separation of the samples in the two groups along the first LV: control samples at negative values (circles) and treated samples at positive ones (squares). The loading plot identifies 21 spots at negative loadings on LV_1 (down-regulated after

the treatment) and 27 at positive loadings on the same LV (up-regulated after the treatment).

Fig. 7(d) represents the coefficients of the original variables calculated for the first LV included in the model: as for the previous case, the coefficients are represented on a virtual map on a colour scale from dark blue (large negative coefficients) to dark red (large positive values). Variables with large negative coefficients correspond to spots down-regulated after the treatment, while those characterised by positive coefficients are up-regulated after the treatment.

The performances of the model appear quite good both in fitting and prediction, according to the R^2 values (table 3b); however, the model performs worse than the one built with variable selection, particularly in regard to the prediction ability. This is particularly visible when looking at the RMSE values in fitting and prediction, showing that the error committed by the model based on PDQuest analysis is about 5 times larger than that committed by the model based on variable selection. The model based on multivariate analysis is thus more reliable above all for what regards prediction. This is also confirmed by the representation of calculated vs. measured responses in fitting and prediction (Fig. 8(d)), showing large variations along the y-axis with respect to BE-PLS-DA model (Fig. 8(c)).

Table 5 reports the spots identified as relevant by BE-PLS-DA and PLS-DA based on PDQuest analysis. BE-PLS-DA identifies 89.58% of the spot identified also by PDQuest (43 of 48). Barely 10.42% (5 of 48) are identified by PDQuest only, while BE-PLS-DA identifies 103 spots more. The spots are listed according to their decreasing positive or negative coefficient.

2.4.5 CONCLUSIONS

PLS-DA was applied coupled to variable selection procedures to identify the differentially expressed proteins in two case-studies: nuclei and lysates from the human colon cancer cell line HCT116 treated by the HDAC inhibitor RC307. The results obtained were compared to those provided by a classical PDQuest differential analysis based on individual Student t-tests (α=0.05).

The multivariate procedure applied with variable selection provided the best results for what regards both the identification of a higher number of significant spots and the achievement of a model characterised by the best predictive ability. The results proved that PDQuest differential analysis identifies a subset of significant spots but this subset does not take into account the multivariate structure of the data and the relationships among the spot volumes. Anyway, both approaches (univariate and multivariate methods) have a role to play in proteomic data analysis. Multivariate tools provide better results above all for what regards the predictive performances since they are able to account for the relationships existing among the variables (spots) and their correlation. The univariate method (such as Student's t-test performed by PDQuest software) instead is actually the simplest to interpret conceptually and most commonly used.

Table 7. Lysates dataset: BE-PLS-DA coefficients of the model containing the first LV.

Down-Regulated										Up-Regulated									
SSP	**Coeff.**	**Both**	**BE**	**PDQuest**	**SSP**	**Coeff.**	**Both**	**BE**	**PDQuest**	**SSP**	**Coeff.**	**Both**	**BE**	**PDQuest**	**SSP**	**Coeff.**	**Both**	**BE**	**PDQuest**
SSP8201	-0.0132	X			SSP1403	-0.00936	X			SSP6902	0.0128	X			SSP2202	0.00764		X	
SSP7602	-0.0123		X		SSP2304	-0.00932		X		SSP6903	0.0128	X			SSP903	0.00758		X	
SSP4308	-0.0123		X		SSP7605	-0.00930		X		SSP3502	0.0124		X		SSP7009	0.00721		X	
SSP1804	-0.0122	X			SSP6503	-0.00928		X		SSP6901	0.0123	X			SSP9305	0.00718		X	
SSP9501	-0.0122		X		SSP9202	-0.00917		X		SSP2813	0.0119	X			SSP7505	0.00690		X	
SSP4205	-0.0120	X			SSP6310	-0.00913		X		SSP3604	0.0117	X			SSP106	0.00677		X	
SSP5204	-0.0119		X		SSP5301	-0.00911		X		SSP1704	0.0115	X			SSP9306	0.00672		X	
SSP8101	-0.0118		X		SSP2103	-0.00907		X		SSP2205	0.0112	X			SSP4107	0.00636		X	
SSP4401	-0.0116	X			SSP2806	-0.00906		X		SSP2403	0.0111	X			SSP2401	0.00628		X	
SSP6401	-0.0115		X		SSP7406	-0.00902		X		SSP905	0.0111	X			SSP1813	0.00620		X	
SSP7305	-0.0114		X		SSP2303	-0.00897		X		SSP1703	0.0110	X			SSP4105	0.00591		X	
SSP7204	-0.0111		X		SSP5205	-0.00888		X		SSP5506	0.0110		X		SSP9004	0.00577		X	
SSP3002	-0.0111		X		SSP6308	-0.00886		X		SSP1	0.0109		X		SSP5804				X
SSP8003	-0.0110	X			SSP4206	-0.00881		X		SSP706	0.0109		X		SSP604				X
SSP8006	-0.0110		X		SSP4801	-0.00874		X		SSP4202	0.0107	X							
SSP8005	-0.0109		X		SSP1202	-0.00870		X		SSP6603	0.0107	X							
SSP5002	-0.0108	X			SSP9702	-0.00849		X		SSP5401	0.0106	X							
SSP9402	-0.0108		X		SSP4304	-0.00848		X		SSP108	0.0105	X							
SSP8707	-0.0108	X			SSP2505	-0.00848	X			SSP6904	0.0104	X							
SSP9303	-0.0107		X		SSP5102	-0.00847	X			SSP201	0.0104	X							
SSP4108	-0.0106		X		SSP3810	-0.00843		X		SSP7012	0.0102	X							
SSP7503	-0.0106	X			SSP7609	-0.00835		X		SSP5807	0.00996	X							
SSP8301	-0.0105	X			SSP6609	-0.00832		X		SSP606	0.00994	X							
SSP1205	-0.0105	X			SSP2404	-0.00829		X		SSP6802	0.00994		X						
SSP7001	-0.0104	X			SSP8703	-0.00816		X		SSP4208	0.00991		X						
SSP8403	-0.0104		X		SSP1508	-0.00813		X		SSP9304	0.00976		X						
SSP5003	-0.0104		X		SSP7014	-0.00809		X		SSP802	0.00961	X							
SSP8305	-0.0103		X		SSP3802	-0.00809		X		SSP7501	0.00953	X							
SSP3206	-0.0102		X		SSP5707	-0.00794		X		SSP7102	0.00945		X						
SSP5006	-0.0101		X		SSP8404	-0.00788		X		SSP502	0.00944		X						
SSP3209	-0.0101		X		SSP7205	-0.00782		X		SSP707	0.00938	X							
SSP4806	-0.0101		X		SSP8504	-0.00781		X		SSP5803	0.00937		X						

Table 6 *(continued)*

SSP4509	-0.0101	X			SSP4702	-0.00751		X		SSP406	0.00927		X					
SSP5306	-0.00998		X		SSP1810	-0.00728		X		SSP705	0.00924		X					
SSP3210	-0.00998		X		SSP7005	-0.00717		X		SSP7402	0.00923	X						
SSP5103	-0.00995		X		SSP2808	-0.00712		X		SSP3103	0.00922		X					
SSP7504	-0.00992	X			SSP8608	-0.00699		X		SSP6303	0.00896		X					
SSP4203	-0.00992	X			SSP2207	-0.00695		X		SSP607	0.00894		X					
SSP7601	-0.00964		X		SSP4703	-0.00686		X		SSP1404	0.00888		X					
SSP8002	-0.00963		X		SSP2302	-0.00681		X		SSP4602	0.00883	X						
SSP7105	-0.00961		X		SSP9505	-0.00679		X		SSP7301	0.00821		X					
SSP803	-0.00956		X		SSP8604	-0.00674		X		SSP5104	0.00818		X					
SSP7707	-0.00950		X		SSP703				X	SSP1302	0.00812		X					
SSP9301	-0.00949	X			SSP4309				X	SSP5704	0.00791		X					
SSP8103	-0.00938		X		SSP4406				X	SSP7109	0.00791		X					
SSP8601	-0.00936		X							SSP9204	0.00768	X						

2.5 REFERENCES

[1] A. Jemal, A. Thomas, T. Murray, M. Thun, CA Cancer J. Clin. 52 (2002) 23-7.

[2] Washington. World Cancer Research Fund and American Institute for Cancer Research, WCRF and AICR Food, nutrition and prevention of cancer: a global perspective. AIoC Research (ed.), 1997.

[3] E. Giovannucci, M.E. Martinez, Tobacco, J. Natl. Cancer Inst. 88 (1996) 1717-730.

[4] R. Labianca, G. Beretta, G. Gatta, F. de Braud, J.Wils, Crit. Rev. Oncol./Hematol. 51 (2004) 145-70.

[5] A. de la Chapelle, Nat. Rev., Cancer 4 (2004) 769-80.

[6] E.R. Fearon, B. Vogelstein, Cell 61 (1990) 759-767.

[7] T. Reya, H. Clevers, Nature 434 (2005) 843-850.

[8] C.A. Purdie, J. O'Grady, J. Piris, A.H. Wyllie, C.C. Bird, Am. J. Pathol. 138 (1991) 807-813.

[9] C. Lengauer, K.W. Kinzler, B. Vogelstein, Nature 396 (1998) 643-649.

[10] Z. Wang, J.M. Cummins, D. Shen, D.P. Cahill, P.V. Jallepalli, T.L. Wang, D. Parsons, G. Traverso, M. Awad, N. Silliman, J. Ptak, S. Szabo, J.K.V. Willson, S.D. Markowitz, M.L. Goldberg, R. Karess, K.W. Kinzler, B. Vogelstein, V.E. Velculescu, C. Lengauer, Cancer Res. 64 (2004) 2998-3001.

[11] H.T. Lynch, A. de la Chapelle, N Engl J Med. 348 (2003) 919-932.

[12] D.H. Robbins, S.H. Itzkowitz, Med Clin North Am. 86 (2002) 1467-1495.

[13] P.M. Harrison, A. Kumar, N. Lang, M. Snyder, M. Gerstein, Nucleic Acids Res. 30 (2002)1083-1090.

[14] C. Pasquali, I. Fialka, L.A. Huber, J Chromatogr B Biomed Sci Appl. 722 (1999) 89-102

[15] L.A. Huber, Nat Rev Mol Cell Biol. 4 (2003) 74–80.

[16] S. Nagata, Annu. Rev. Immunol. 23 (2005) 853–875.

[17] J. Gotzmann, S. Vlcek, R. Foisner, J. Cell Sci. 113 (2000) 3769– 3780.

[18] A.J.M. De Ruijter, A.H. van Gennip, H.N. Caron, S. Kemp, B.P. van Kuilenburg, Biochem. J. 370 (2003) 737-749.

[19] R.W. Johnstone, J.D. Licht, Cancer Cell 4 (2003) 13-18.

[20] R. Brown, G. Strathdee, Trends Mol. Med. 8 (2002) 43-48.

[21] R.W. Johnstone, Nat. Rev. Drug Discov. 12, (2002) 294-300.

[22] F. McLaughlin, N.B. La Thangue, Biochem. Pharmacol. 68, (2004) 1139-1144.

[23] D. Cecconi, A. Scarpa, M. Donadelli, M. Palmieri, M. Hamdan, H. Astner, P.G. Righetti, Electrophoresis 24 (2003) 1871-1878.

[24] D. Cecconi, M. Donadelli, A. Scarpa, A. Milli, M. Palmieri, M. Hamdan, J. Rappsilber, L.B. Areces, P.G. Righetti, J. Proteome Research 4 (2005) 1900-1916.

[25] D. Cecconi, M. Donadelli, S. Rinalducci, L. Zolla, A. Scarpa, M. Palmieri, P.G. Righetti, Proteomics 7 (2007) 1644-1653.

[26] J-H. Park, Y. Jung, T.Y. Kim, S.G. Kim, H-S. Jong, J.W. Lee, D.K. Kim, J.S. Lee, N.K. Kim, T.Y. Kim, Y.J. Bang, Clin. Cancer Res. 10 (2004) 5271-5281.

[27] T. Beckers, C. Burkhardt, H. Wieland, P. Gimmnich, T. Ciossek, T. Maier, K. Sanders, Int. J. Cancer 121 (2007) 1138-1148.

[28] P. Atadja, L. Gao, P. Kwon, N. Trogani, H. Walker, M. Hsu, L. Yeleswarapu, N. Chandramouli, L. Perez, R. Versace, A. Wu, L. Sambucetti, P. Lassota, D. Cohen, K. Bair, A. Wood, S. Remiszewski, Cancer Res. 64 (2004) 689-695.

[29] D. Cecconi, S. Mion, H. Astner, E. Domenici, P.G. Righetti, L. Carboni, Brain Research 1135 (2007) 41-51.

[30] F. Al-Shahrour, P. Minguez, J. Tárraga, D. Montaner, E. Alloza, J.M. Vaquerizas, L. Conde, C. Blaschke, J. Vera, J. Dopazo, Nucleic Acids Research 34 (2006) W472-W476.

[31] K. Endo, S. Kohnoe, E. Tsujita, A. Watanabe, H. Nakashima, H. Baba, Y. Maehara, Anticancer Res. 25 (2005) 3117-21.

[32] C. Greco, R. Vona, M. Cosimelli, P. Matarrese, E. Straface, P. Scordati, D. Giannarelli, V. Casale, D. Assisi, M. Mottolese, A. Moles, W. Malorni, Glycobiology 14 (2004) 783-92.

[33] S. Akahani, P. Nangia-Makker, H. Inohara, H.R. Kim, A. Raz, Cancer Res. 57 (1997) 5272-6.

[34] R.Y. Yang, D.K. Hsu, F.T. Liu, Proc Natl Acad Sci U S A. 93 (1996) 6737-42.

[35] D.M. Townsend, K.D. Tew, Oncogene 22 (2003) 7369–75.

[36] M.N. Ebert, G. Beyer-Sehlmeyer, U.M. Liegibel, T. Kautenburger, T.W. Becker, B.L. Pool-Zobel, Nutr Cancer 41 (2001) 156–64.

[37] B.L. Pool-Zobel, V. Selvaraju, J. Sauer, T. Kautenburger, J. Kiefer, K.K. Richter, M. Soom, S. Wölfl, Carcinogenesis 6 (2005) 1064–76.

[38] W.H. Lee, R.A. Morton, J.I. Epstein, J.D. Brooks, P.A. Campbell, G.S. Bova, W.S. Hsieh, W.B. Isaacs, W.G. Nelson, Proc Natl Acad Sci USA 91 (1994) 11733–11737.

[39] C.A. Neumann, Q. Fang, Curr Opin Pharmacol. 7 (2007) 375-80.

[40] R.A. Egler, E. Fernandes, K. Rothermund, S. Sereika, N. de Souza-Pinto, P. Jaruga, M. Dizdaroglu, E.V. Prochownik, Oncogene 24 (2005) 8038-50.

[41] I. Hoshino, H. Matsubara, N. Hanari, M. Mori, T. Nishimori, Y. Yoneyama, Y. Akutsu, H. Sakata, K. Matsushita, N. Seki, T. Ochiai, Clin Cancer Res. 11 (2005) 7945-52.

[42] S.B. Cullinan, L. Whitesell, Semin Oncol. 33 (2006) 457-65.

[43] C. Dai, L. Whitesell, Future Oncol. 1 (2005) 529-40.

[44] S.Y. Sharp, C. Prodromou, K. Boxall, M.V. Powers, J.L. Holmes, G. Box, T.P. Matthews, K.M. Cheung, A. Kalusa, K. James, A. Hayes, A. Hardcastle, B. Dymock, P.A. Brough, X. Barril, J.E. Cansfield, L. Wright, A. Surgenor, N. Foloppe, R.E. Hubbard, W. Aherne, L. Pearl, K. Jones, E. McDonald, F. Raynaud, S. Eccles, M. Drysdale, P. Workman, Mol Cancer Ther. 6 (2007) 1198-211.

[45] J.I. Johnsen, O.N. Aurelio, Z. Kwaja, G.E. Jörgensen, N.S. Pellegata, R. Plattner, E.J. Stanbridge, J.F. Cajot, Int J Cancer. 88 (2000) 685-91.

[46] R. Wang, K. Dong, F. Lin, X. Wang, P. Gao, S.H. Wei, S.Y. Cheng, H.Z. Zhang, Mol Med. 13 (2007) 567-75.

[47] S. Singer, V. Ehemann, A. Brauckhoff, M. Keith, S. Vreden, P. Schirmacher, K. Breuhahn, Hepatology 46 (2007) 759-68.

[48] K.C. Sedoris, S.D. Thomas, D.M. Miller, Biochemistry 46 (2007) 8659-68.

[49] K. Ejeskär, C. Krona, H. Carén, F. Zaibak, L. Li, T. Martinsson, P.A. Ioannou, BMC Cancer 5 (2005) 161.

[50] S. Fukuda, D.W. Wu, K. Stark, L.M. Pelus, Biochem Biophys Res Commun. 292 (2002) 593-600.

[51] S. Fukuda, L.M. Pelus, Cell Mol Life Sci. 62 (2005) 1526-7.

[52] T. Fillies, R. Werkmeister, J. Packeisen, B. Brandt, P. Morin, D. Weingart, U. Joos, H. Buerger, BMC Cancer. 6 (2006) 10.

[53] H.K. Parekh, H. Simpkins, Cancer Res. 55 (1995) 5203-6.

[54] B. Schutte, M. Henfling, F.C. Ramaekers, Apoptosis 11 (2006) 1561-72.

[55] S. Gilbert, A. Loranger, N. Marceau, Mol Cell Biol 24 (2004) 7072–7081.

[56] M.D. Galigniana, J.L. Scruggs, J. Herrington, M.J. Welsh, C. Carter-Su, P.R. Housley, W.B. Pratt, Mol. Endocrinol. 12 (1998) 1903-1913.

[57] P.A. Clarke, I. Hostein, U. Banerji, F.D. Stefano, A. Maloney, M. Walton, I. Judson, P. Workman, Oncogene 19 (2000) 4125-33.

[58] N.L. Anderson, J.P. Hofmann, A. Gemmell, J. Taylor, Clin. Chem. 30 (1984) 2031-2036.

[59] P. Tarroux, P. Vincens, T. Rabilloud, Electrophoresis 8 (1987) 187-199.

[60] I. Gadea, G. Ayala, M.T. Diago, A. Cunat, J. Garcia de Lomas, Clin. Diagn. Lab. Immunol. 7 (2000) 549-552.

[61] H. Kovarova, M. Hajduch, G. Korinkova, P. Halada, S. Krupickova, A. Gouldsworthy, N. Zhelev, M. Strnad, Electrophoresis 21 (2000) 3757-3764.

[62] Y. Iwadate, T. Sakaida, T. Hiwasa, Y. Nagai, H. Ishikura, M. Takiguchi, A. Yamaura, Cancer Res. 64 (2004) 2496-2501.

[63] R.A. Amin, A.E. Vickers, F. Sistare, K.L. Thompson, R.J. Roman, M. Lawton, J. Kramer, H.K. Hamadeh, J. Collins, S. Grissom, L. Bennett, C.J. Tucker, S. Wild, C. Kind, V. Oreffo, J.W. Davis, S. Curtiss, J.M. Naciff, M. Cunningham, R. Tennant, J. Stevens, B. Car, T.A. Bertram, C.A. Afsharil Environ. Health Perspect. 112 (2004) 465-479.

[64] W.H.M. Heijne, R.H. Stierum, M. Slijper, P.J. van Bladeren, B. van Ommen, Biochem. Pharmacol. 65 (2003) 857-875.

[65] K.C.M. Verhoeckx, S. Bijlsma, E.M. de Groene, R.F. Witkamp, J. van der Greef, R.J.T. Rodenburg, Proteomics 4 (2004) 1014-1028.

[66] D. Magdic, D. Horvat, Z. Jurkovic, R. Sudar, K. Kurtanjek, Food Technol. Biotechnol. 40 (2002) 331-341.

[67] Marengo E, Robotti E, Righetti PG, Campostrini N, Pascali J, Ponzoni M (2004) Clin. Chim. Acta 345: 55-67.

[68] E. Marengo, E. Robotti, D. Cecconi, A. Scarpa, P.G. Righetti, Anal. Bioanal. Chem. 379 (2004) 992-1003.

[69] E. Marengo, E. Robotti, F.Antonucci, D. Cecconi, N. Campostrini, P.G. Righetti, Proteomics 5 (2005) 654-666.

[70] E. Marengo, E. Robotti, M. Bobba, M.C. Liparota, F. Antonucci, C. Rustichelli, A. Zamò, M. Chilosi, M. Hamdan, P.G. Righetti, Electrophoresis 27 (2006) 484-494.

[71] K.S. Lilley, P. Cupree, J. Exp. Botany 57 (2006) 1493-1499.

[72] P. Pir, B. Kirdar, A. Hayes, Z.Y. Onsan, K.O. Ulgen, S.G. Oliver, BMC Bioinformatics 7 (2006) art. No. 203.

[73] K.C.M. Verhoeckx, M. Gaspari, S. Bijlsma, J. van der Greef, R.F. Witkamp, R.P. Doornbos, R.J.T. Rodenburg, J. Proteome Res. 4 (2005) 2015-2023.

[74] F. Jessen, R. Lametsch, E. Bendixen, I.V.H. Kjaersgard, B.M. Jorgensen, Proteomics 2 (2002) 32-35.

[75] T.G. Kleno, L.R. Leonardsen, H.O. Kjeldal, S.M. Laursen, O.N. Jensen, D. Baunsgaard, Proteomics 4 (2004) 868-880.

[76] I.V.H. Kjaersgard, M.R. Norrelykke, F. Jessen, Proteomics 6 (2006) 1606-1618.

[77] N.A. Karp, J.L. Griffin, K.S. Lilley, Proteomics 5 (2005) 81-90.

[78] B. Norden, P. Broberg, C. Lindberg, A. Plymoth, Chem. Biodiv. 2 (2005) 1487-1494.

[79] D.L. Massart, B.G.M. Vandeginste, S.M. Deming, Y. Michotte, L. Kaufman, Chemometrics: a textbook, Elsevier, Amsterdam, 1988.

[80] B.G.M. Vandeginste, D.L. Massart, L.M.C. Buydens, S. De Jong, P.J. Lewi, J. Smeyers-Verbeke Handbook of Chemometrics and Qualimetrics: Part B, Elsevier, Amsterdam, 1998.

[81] R.A. Eisenbeis (Ed.), Discriminant Analysis and Classification Procedures: Theory and Applications, Lexington Books, Lexington, 1972.

[82] H. Martens, T. Naes, Multivariate Calibration, Wiley, London,1989.

[83] D. Kleinbaum, L. Kupper, K. Muller, Applied Regression Analysis and Other Multivariate Methods, 2nd ed., Pws-Kent, Boston, 1988.

CHAPTER 3

PROTEOMIC ANALYSIS OF WINE LACTIC ACID BACTERIA

3.1 WINE BACTERIA AND MALOLACTIC FERMENTATION (MLF)

Wine may be considered a particularly stressful environment in which several lactic acid bacteria (LAB), mainly belonging to the *Lactobacillus, Pediococcus*, and *Oenococcus* genera, acetic acid bacteria (AAB), and to a lesser extent species of bacilli, have developed different mechanisms aimed to best perform their biological roles. LAB and AAB are often present on the surface of the grape and can represent significant populations in musts. Most LAB found in wine, including members of *Oenococcus, Lactobacillus, Pediococcus*, and *Leuconostoc*, are microaerophilic and able to grow in the anaerobic environment of fermenting wine. In contrast, wine-related AAB, such as members of *Gluconobacter* or *Gluconacetobacter*, are obligately aerobic and loosely categorized as vinegar bacteria [1]. Both AAB and LAB can produce acetic acid, a potential inhibitor of growth and alcohol production by *Saccharomyces cerevisiae* [2]. Wine LAB are responsible for the malolactic fermentation (MLF), which usually follows the alcoholic fermentation performed by yeast, principally by *Saccharomyces cereviseae*, and can begin only when the bacterial population reaches 10^6 CFU/mL [2]. MLF in wine is a biological process of wine deacidification in which the dicarboxylic l-malic acid (malate) is converted to the monocarboxylic l-lactic acid (lactate) and carbon dioxide. *Oenococcus oeni* is the preferred species used to conduct MLF due to its acid tolerance and flavor profile produced. This secondary fermentation increases the pH, improves the taste and the sensory quality of wine. Moreover, MLF stabilizes the wine microbiologically by consumption of residual carbon nutrients. In addition to its occurrence in wine, MLF occurs in other fermented beverages, such as cider [1, 3]. It is widely accepted that the role of MLF is more than just a deacidification process, although deacidification via MLF is still a primary objective in wine fermentation, especially in cool climate regions. The complexity and diversity of the metabolic activity of wine LAB suggest that MLF may affect wine quality both positively and negatively.

Although some wine LAB, such as *O. oeni* and *L. plantarum*, have been extensively studied for their predominant role during MLF and basic knowledge has increased in recent years, many gaps still exist. The most evident is relative to the natural adaptation of wine LAB to such a harsh medium [2, 4].

Numerous reviews on MLF have been published, covering the history of research on MLF, taxonomy, ecology, genetics, and physiology of wine LAB and the oenological significance of MLF [2, 5, 6].

3.1.1 STRESS AND STRESSORS IN WINE

Stress can be defined as a change in the genome, the proteome, or the environment, producing a decrease in the growth rate or survival. Any form of life may experience stress if the conditions, to which they are accostumed, undergo rapid changes. Stress responses are of particular importance for microorganisms, because their habitats are subject to continual changes (such parameters as temperature, osmotic pressure, and substrate availability are far from being constant). Stressors, or stress factors, may have a chemical, physical, or biological nature. Although some stresses (e.g., high/low temperature, high/low pH value or ethanol concentration) are induced by the environment, natural stresses such as acidity and, sometimes, starvation are generated by cell growth itself [7, 8]. The organism itself may be the source of stress factors. For example, many bacteria produce reactive oxygen species (ROSs) that may cause oxidation of macromolecules, protein denaturation and nucleic acids degradation.

In wine various stress factors have a negative impact on LAB growth. Wine is a medium with low pH, usually 3–3.5, because of the presence of several organic acids. It also contains sulphur dioxide (50–100 mg/L) added as antimicrobial compound and MFL has to occur at low temperatures, usually below 18°C. Moreover, during alcoholic fermentation, grapes must have a composition which is modified by yeast metabolism. Basically, *S. cerevisiae* produces ethanol to levels as high as 12–13% (v/v) and sometimes sulphur dioxide up to 100 mg/L for some strains. In addition, other yeast metabolites such as aromatic compounds like phenylethanol and proteins have been identified as inhibitors for wine LAB, specially *O. oeni* [9, 3]. Moreover, stress compounds in wine such as fatty acids and copper have an inhibitory effect on cell

growth and MLF duration, an effect that seems to be correlated to an inhibition of ATPase activity. The harsh environment resulting from both accumulation of the lactic acid produced during carbohydrate catabolism and generated stress factors, limits the growth and the survival of wine LAB and may facilitate the selection of particular species contributing indirectly to the final taste of wine. However, also in the stressful environment generated, wine LAB such as *L. plantarum*, are able to undergoing biochemical process such as MFL.

Although *L. plantarum* is capable of malolactic fermentation, it usually contributes to production of undesirable products such as biogenic amine and precursors of ethyl carbammate [1, 2, 10] and is therefore of general concern because of its spoilage nature. However, *L. plantarum* is frequently isolated from red wine undergoing MFL and sterilised with sulphite [11]. Therefore, either the “good” *O. oeni* or the “bad” *L. plantarum* have developed several mechanisms to escape or to tolerate wine conditions and their ability to survive strictly depends from the stress response machinery developed.

3.2 EFFECT OF TANNIC ACID ON *LACTOBACILLUS PLANTARUM* WINE STRAIN DURING STARVATION: A PROTEOMIC STUDY

3.2.1 INTRODUCTION

Tannins are phenols widespread in the plant kingdom and they are mainly classified into hydrolysable and condensed tannins. Vegetable tannins are present in soil, in foodstuffs and, as consequences, on the alimentary tracts of humans and animals [12]. Bacteria, harboring these different environments, can interact with tannins which affect the bacteria growth [13]. For example, it has been demonstrated that tannins interfere on physiology of LAB [14-16], and that tannic acid (TA), a hydrolysable tannin, has an inhibitory effect on the growth of intestinal bacteria [14]. Accordingly, we have recently demonstrated, by a proteomic approach, that TA inhibits the growth of *Lactobacillus hilgardii* (a wine spoilage bacterium) [17].

Some bacteria are able to degrade tannins by producing tannin acyl hydrolase (E.C. 3.1.1.20), commonly referred as tannase [18, 19]. Previous investigations on tannin-degrading bacteria regarded soil species [20, 21]; afterward tannase activity has also been detected in LAB cells [22]. Several tannin-degrading LABs, such as *L. plantarum*, *L. paraplantarum* and *L. pentosus*, have been isolated from humans feces and from fermented foods [23, 24], while only *L. plantarum* has been detected among the tannase-positive oenological LAB species [25]. *L. plantarum* is a versatile Gram-positive fermentative bacterium that can be found in a range of habitats, including dairy, meat, and many plant fermentations. Thus, the *L. plantarum* strains are able to adapt to different environmental conditions. In fact, it has been suggested that tannase activity might confers an ecological advantage to bacteria [22, 26]. Moreover, it has been demonstrated that different bacteria strains [20, 27], and in particular *L. plantarum* [28, 29], are able to use TA as sole carbon source by inducible production of tannase.

In the last years, tannase has been the subject of a lot of studies due to its commercial importance and complexity as catalytic molecule [19, 30]. In particular, interest on the *L. plantarum* tannase activity has been addressed for aspects of dietary of tannins-rich foods and for its application in food and beverage industrial processes [14]. Nevertheless the information concerning the effects of tannins on the physiology of

tannase-positive LAB is as yet scarce. In order to elucidate the interaction between tannins and bacteria further investigations on the response of tannase-positive LAB to tannins are recommended.

In the present study (carried out in collaboration with Dr. Zapparoli laboratoty, Department of Biotechnology, University of Verona), the effects of TA on a tannase-positive wine strain VP08 of *L. plantarum* were examined. Tannase activity and growth kinetics were quantified in cells cultivated in semi-synthetic medium containing TA or glucose as carbon source.

The growth kinetics analysis was performed using different TA concentrations: 100 and 250 mg/L (as enological concentrations of hydrolysable tannins that can be present in wine [31]), and 500 mg/L (as an excess amount). Glucose was used at 2 g/L in order to reproduce the residual sugar concentration present in wine after alcoholic fermentation.

Then, a comparative 2D-PAGE analysis of total proteins extracted from cells in the stationary phase (grown in 250 mg/L TA or 2 g/L glucose) was carried out. We show that tannic acid slows down the proliferation and prolongs the viability of *L. plantarum* cells; moreover TA regulates the expression of proteins of particular interest involved in glycolysis, amino acid metabolism, translation and protein folding. The data provided insights into the possible mechanisms involved in the response of *L. plantarum* cells to tannins during starvation.

3.2.2 MATERIAL AND METHODS

3.2.2.1 Bacterial Strain And Culture Conditions

In this study the strain VP08 belonging to *L. plantarum* was used. This strain was isolated from red wine containing oenological tannins which were added (as winemaking practice) during the beginning of alcoholic fermentation. The strain VP08 was routinely cultivated at 28°C in MRS broth (Fluka, Seelze, Germany) pH 4.8.

3.2.2.2 Tannase Assay

The enzymatic assay was carried out on whole cells and on the supernatant obtained by centrifugation of cell culture grown in MRS and afterwards in semi-synthetic medium

(composed of casamino acid 5 g/L, yeast extract 4 g/L, $MnSO_4$ 0.025 g/L, $MgSO_4$ 0.125 g/L, KCl 0.425 g/L, KH_2PO_4 0.550 mM, $CaCl_2$ 0.125 mM, pH 4.5) added with 250 mg/L TA or 2g/L glucose. The pellet was washed twice in 2 mL of 33 mM NaH_2PO_4 buffer, pH 5.0, before using the cells for the assay. A total volume of 1 mL of the supernatant was filtered (0.22 μm pore size membranes; Millipore, Bedford, Massachusetts, USA) and concentrated 20 fold using Centricon (cut-off 5000 Da, Millipore). Tannase activity of the whole cells and supernatant was quantified by a colorimetric method as described by Nishitani and Osawa [32]. Briefly, the sample preparation was as follows: the pellet of whole cells, grown on MRS, was suspended in 1 mL of substrate medium (pH 5.0) containing 33 mM NaH_2PO_4 and 10 mM methyl gallate at final concentration to prepare a suspension whose absorbance at 600 nm was adjusted to 1 (=ca. 1.5 x 10^8 CFU/mL); while the concentrated supernatant (50 μL) was suspended in 950 μl of substrate medium at a final concentration of 5% v/v. The two suspensions were incubated aerobically at 30°C for 2h, centrifuged (8000xg, 5 min), and the absorbance of supernatants was read at 450 nm.

Commercial tannase of *Aspergillus ficuum* (Sigma, St. Louis, MO, USA) was used for standard calibration. The activity was expressed in mU. By definition 1 mU of tannase hydrolyzes 1 nmol of TA *per* min at pH 5.0 at 30°C. The experiment was carried out in triplicate using independent samples. Statistical treatment of data was carried out by t-test and significant differences were at $p<0.05$.

3.2.2.3 Growth Kinetics

The growth kinetic of the *L. plantarum* VP08 strain was tested in semi-synthetic medium, using TA as carbon source at the concentration of 100, 250 and 500 mg/L (TA 100, TA 250 and TA 500) or 2 g/L glucose.

An amount of 1 mL of overnight culture cells grown in MRS medium pH 4.8 was harvested, washed in 1 mL of NaCl 0.9%, and inoculated at the concentration of 2 x 10^6 cell/mL in 10 mL of semi-synthetic medium contained in glass sealed tubes. The incubation was carried out at 28°C without agitation. The growth was estimated as CFU/mL by plate count on MRS agar pH 4.8, after the incubation of the plates

aerobically for three days at 28°C. The growing test was carried out in triplicate using independent samples.

3.2.2.4 Total Protein Extraction

Preliminary results using different concentrations of TA and glucose allowed us to observe when cell cultures entered in the stationary phase and how long it lasted.

Thus, we have chosen 250 mg/L TA and 2 g/L glucose, and 6 days from the inoculums, as a good compromise to obtain the cells at the same physiological state (i.e. just before the reduction of viability) and at the same concentration during the stationary phase. Thus, a total amount of 3×10^9 cells grown in 250 mg/L TA or in 2 g/L glucose were harvested after 6 days of incubation at 28°C. The cell pellets were washed twice with 0.9% NaCl and cells were lysed, by vortexing at maximum speed for 30 min at 4°C, in 5 mL of solubilization buffer: 7 M urea (Sigma), 2 M thiourea (Sigma), 20 mM Tris (Sigma), 5 mM tributylphosphine (Sigma), 1% ampholyte pH 3-10 (Fluka, Buchs SG Switzerland), 3% CHAPS (Sigma), 1x protease inhibitor cocktail tablet (Complete, Mini; Roche, Basel, Switzerland), added with 0.1 g micro-glass beads. The lysate were then centrifuged for 10 min at 10,000 x g at 4°C to remove the nucleic acids (complexed with ampholytes), as well as cellular debris, cell membranes, and unsolubilized compounds. The protein concentration was evaluated with the DC Protein assay (Bio-Rad, Hercules, CA) based on the Lowry method. The calibration curve was obtained by using known concentrations of BSA dissolved in distilled water.

3.2.2.5 2D-PAGE Analysis Of Total Protein Extract

Two-dimensional gel electrophoresis was performed as described in Section 2.3.2.5.

3.2.2.6 Protein Pattern Differential Analysis

Gel scanning and image analysis of the two separated groups (TA and glucose-grown cells) were performed as decribed in Section 2.3.2.6. Data were log transformed and quantitative and a Student's t-test were performed in order to compare the two groups and identify sets of proteins that showed a statistically significant difference with a confidence level of 0.05 and a minimum two fold of variation.

3.2.2.7 In-Gel Digestion

Spots showing a statistically significant differential expression were carefully manually cut out from 2-D Sypro Ruby stained gels and subjected to in-gel trypsin digestion according to Shevchenko *et al.* [33] with minor modifications. For a further description see also Section 2.3.2.7.

3.2.2.8 Peptide Sequencing By Nano HPLC-ESI-MS/MS

Peptide mixtures were separated by using a nanoflow-HPLC system (Ultimate; Switchos; Famos; LC Packings, Amsterdam, The Netherlands). A sample volume of 10 µL was loaded by the autosampler onto a homemade 2 cm fused silica pre-column (75 µm I.D.; 375 µm O.D.; Reprosil C18-AQ, 3 µm from Ammerbuch-Entringen, DE) at a flow rate of 2 µL/min. Sequential elution of peptides was accomplished by using a flow rate of 200 nl/min and a linear gradient from Solution A (2% acetonitrile; 0.1% formic acid) to 50% of Solution B (98% acetonitrile; 0.1% formic acid) in 40 minutes over the pre-column in-line with a homemade 10-15 cm resolving column (75 µm I.D.; 375 µm O.D.; Reprosil C18-AQ, 3 µm from Ammerbuch-Entringen, Germany).

Peptides were eluted directly into a High Capacity ion Trap (model HCTplus, Bruker-Daltonik, Germany). The capillary voltage was 1.5-2 kV and a dry gas flow rate of 10 L/min was used with a temperature of 230°C. The scan range used was from 300 to 1800 m/z. Protein identification was performed by searching in the National Center for Biotechnology Information non-redundant database (NCBInr) using the Mascot program (http://www.matrixscience.com). The following parameters were adopted for database searches: complete propionamide formation on cysteines and partial oxidation of methionines, peptide Mass Tolerance ± 1.2 Da, Fragment Mass Tolerance ± 0.9 Da, missed cleavages 2. For positive identification, the score of the result of [-10 x Log(P)] had to be over the significance threshold level ($p < 0.05$).

3.2.2.9 Protein Categorization

Gene ontology (GO) lists were downloaded from web tools for the functional analysis of groups of genes in high-throughput experiments along with the use of information on Gene Ontology terms. Each protein was classified with respect to its biological process and molecular function using GO annotation. When no GO annotation was available,

proteins were annotated manually based on literature searches and closely related homologues.

3.2.3 RESULTS

3.2.3.1 Tannase Assay

By a preliminary assay carried out in whole *L. plantarum* cells grown in MRS the strain VP08 displayed tannase activity (data not shown). This activity was quantified in cells cultivated in the medium with TA and glucose as carbon source and in the corresponding supernatants. Cells grown in TA displayed higher activity than those grown in glucose (p= 0.0192). (Fig. 1). In the supernatants tannase activity was absent.

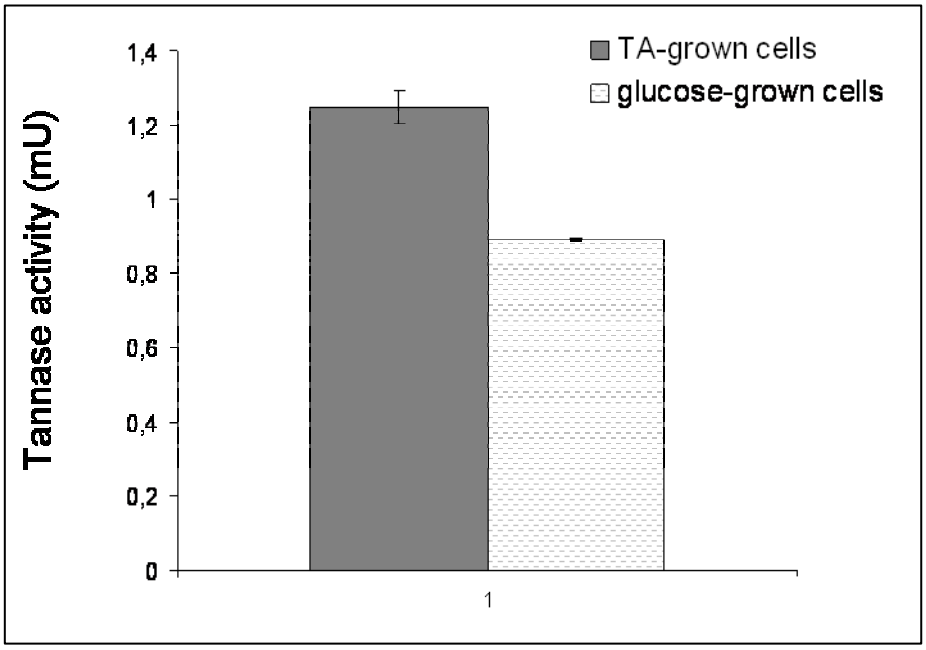

Fig. 1. Tannase activity of *Lactobacillus plantarum* strain VP08 whole cells grown in semi-synthetic medium containing 250 mg/L tannic acid or 2 g/L glucose as carbon source. The tannase activity was quantified in mU *per* 1.5 x 10^8 CFU/mL. Values are the means of three independent experiments.

3.2.3.2 Growth Tests

Fig. 2 shows the growth curves of *L. plantarum* VP08 strain in medium containing 100, 250 and 500 mg/L TA and 2 g/L glucose. The kinetics differed in relationship to the source of carbon and TA concentrations. In glucose, TA 100 and TA 250 cell

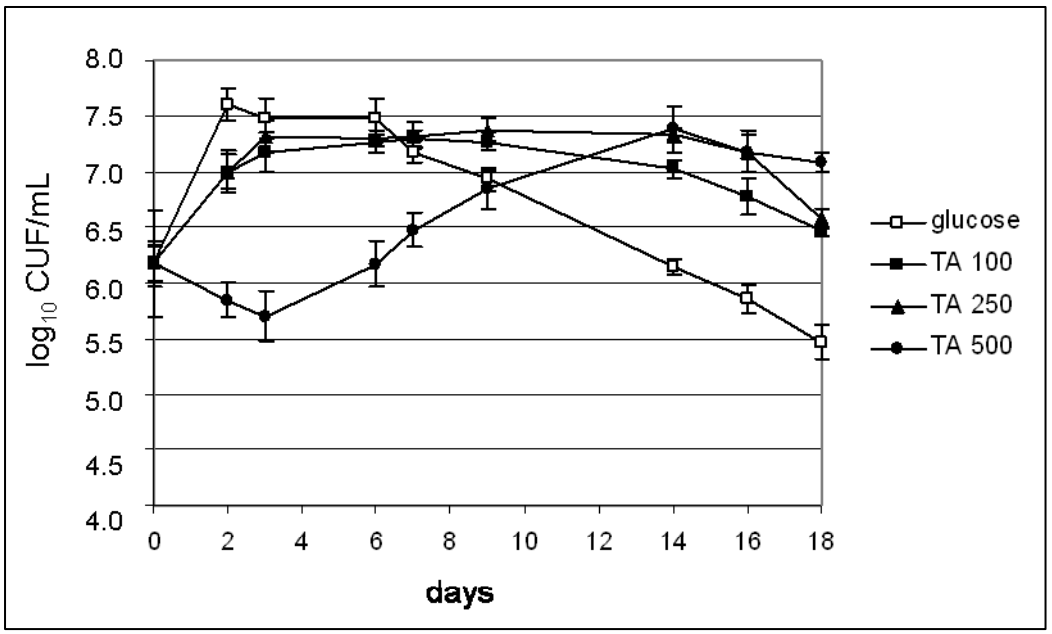

Fig. 2. Growth kinetics of *Lactobacillus plantarum* strain VP08 grown in semi-synthetic medium containing tannic acid (TA) at 100, 250 and 500 mg/L or 2 g/L glucose as carbon source. Values are the means of three independent experiments

concentration increased in two-three days after the inoculation, while in TA 500 cell population underwent partial mortality in the same period. Furthermore, in TA 100 and 250 cell population remained in the stationary phases longer (about 8 days more) than in glucose. After this growth phase a constant rate of mortality was observed. On the contrary, in TA 500 the cell number increased slowly for 12 days reaching the same maximal concentration as that observed in TA 100 and TA 250.

3.2.3.3 *L. plantarum* 2-DE Protein Pattern Analysis

To analyze the effect of TA on the growth of wine strain VP08 of *L. plantarum* the proteome profiles of cells, grown in medium supplemented with 2 g/L glucose or 250 mg/L TA, were evaluated by 2D-PAGE. Protein spots showing a statistically significant differential expression in TA-grown cells were analyzed at 6 days by means of 2-DE in the pH range of 3 to 10. Fig. 3 shows all the 2D maps obtained for glucose and TA-grown cells. By PDQuest analysis we detected protein spots and measured differential protein expression by analyzing the 4 replica of 2D maps obtained for each group. The average number of spots detected was 418 ± 11 and 422 ± 5 for 2D maps of glucose and TA-grown cells respectively. A total of 36 different spots (matched across all the replica 2D maps) were found to be differentially expressed in cells grown in semi-

synthetic medium supplemented with TA: in particular 15 spots were found to be down-regulated and 21 up-regulated.

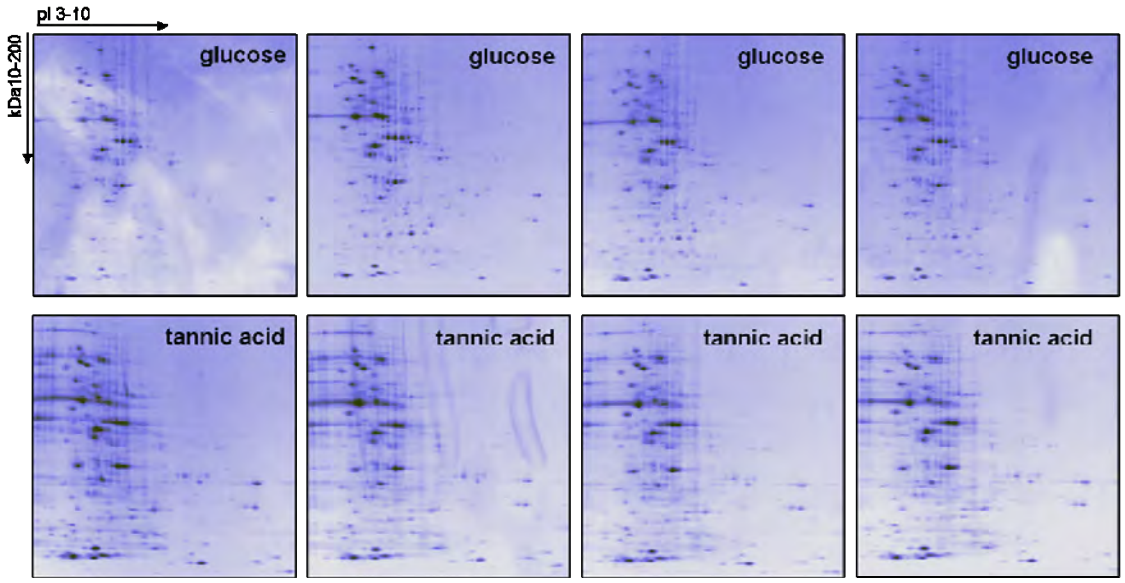

Fig. 3. *Lactobacillus plantarum* 2D maps of glucose and TA-grown cells. The proteins (about 675 μg) were first separated in a linear pH gradient of 3–10, followed by separation in an SDS-PAGE (8-18%) and Sypro Ruby fluorescent staining.

The typical high resolution 2-DE protein pattern obtained from *L. plantarum* protein extract is reported in Fig. 4a, whereas Fig. 4b shows representative, differentially expressed spots ($p<0.05$) between glucose and TA-grown cells. Spots selected from the differential analysis were subjected to RP-HPLC-ESI-MS/MS analysis for protein identification. The unique differentially expressed proteins identified were 18. All the identifications were obtained by using protein sequences of *L. plantarum* WCFS1 for protein matching. In Table 1, the successfully identified proteins corresponding to up- or down-regulated spots are shown, together with the gene names, the spot number, the identification parameters and the indication of their gene ontology (GO) annotation (biological process and molecular function). Most of the spots were identified as proteins involved in glycolysis, amino acid metabolism, translation and protein folding. Interestingly, all the spots identified as glycolytic enzymes (glyceraldehyde 3-phospate dehydrogenase, pyruvate kinase and 6-phosphofructokinase) increased in TA-grown cells. On the contrary those identified as protein related to amino acid metabolism (2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase, cystathionine β-lyase and glicine hydroxymethyltransferase) decreased in TA-grown cells. Fig. 5 shows a pie

chart with the distribution of the identified proteins catalogued according to the biological process in which they are involved.

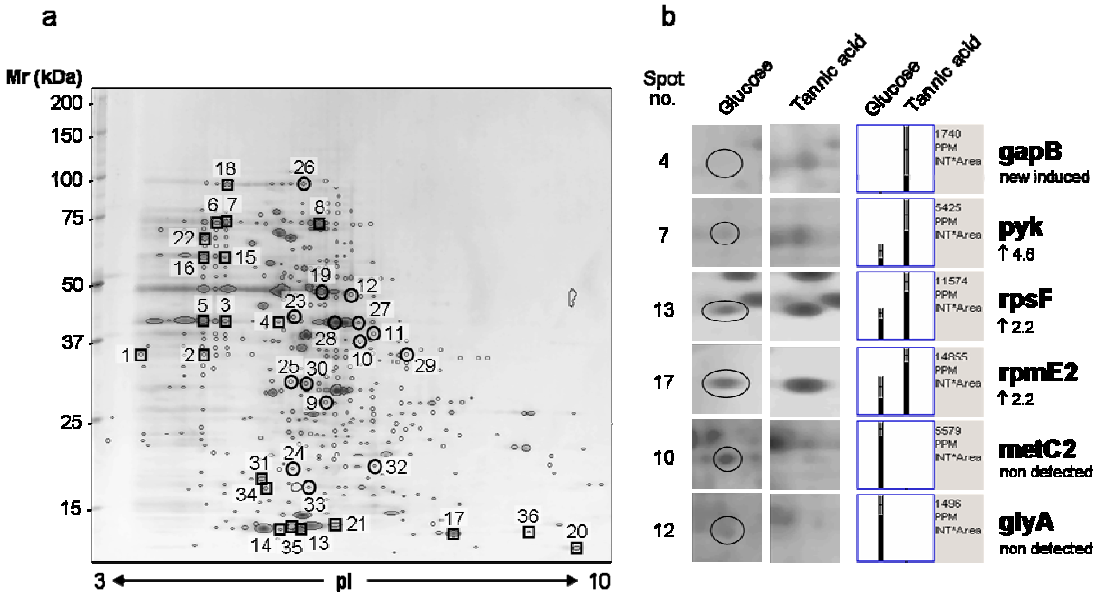

Fig. 4. (a) Representative Sypro Ruby stained 2-D gels of proteins extracted from *Lactobacillus plantarum* cells. The 36 differentially expressed proteins are marked by spot number as reported in Table 1, plus an opened circle (to indicate decreased protein spots) or an opened square (for increased protein spots). (b) PDQuest (version 7.3) output showing representative differentially expressed spots ($p<0.05$) between glucose and tannic acid-grown cells. For each spot, an enlarged region of the respective 2-DE map containing the referred spot (highlighted by an ellipse), as well as, a histogram showing the quantitative variation in the protein expression, are shown. Bars show the average spot quantity (black dash) and standard deviation (vertical white lines).

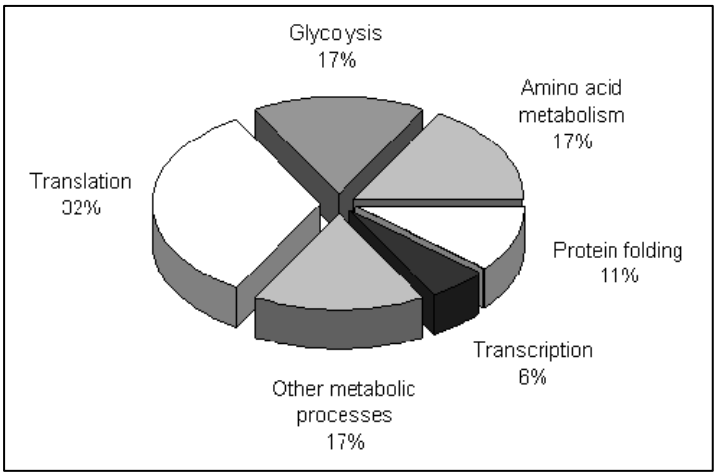

Fig. 5. Distribution of the identified proteins according to the biological function. Assignments were made on the basis of information provided by gene ontology (GO) lists downloaded using the tool FatiGO from Babelomics (http://fatigo.bioinfo.cipf.es/). The absolute number of proteins to which the distribution is referred is: 6 (translation), 3 (glycolysis), 3 (amino acid metabolism), 2 (protein folding), 1 (transcription), and 3 (other metabolic processes).

Table 2. Identification of the differentially expressed proteins in tannic acid-grown cells of *L. plantarum* VP08 strain.

Protein name	Gene name	Spot no. [a] (Fig. 4)	SwissProt TrEMBL acc. #	NCBI acc. #	Mr. (kDa) exp./theor.	pI exp./theor.	No. of peptides identified	Mascot score [b]	Sequence Coverage [c] (%)	Molecular function GO term	Fold of variation after TA [d]
Glycolysis											
6-phosphofructokinase	pfkA	1	Q88VY1	gi|28378549	27.5/34.2	4.0/5.13	3	148	12	Transferase activity	+ 2.7
6-phosphofructokinase	pfkA	2	Q88VY1	gi|28378549	27.5/34.2	4.6/5.13	4	212	15	Transferase activity	new induced
Glyceraldehyde 3-phosphate dehydrogenase	gapB	3	Q88YH6	gi|28377642	33.0/36.6	4.8/5.3	6	312	22	NAD binding	+ 4
Glyceraldehyde 3-phosphate dehydrogenase	gapB	4	Q88YH6	gi|28377642	32.5/36.6	5.3/5.3	9	519	32	NAD binding	new induced
Glyceraldehyde 3-phosphate dehydrogenase	gapB	5	Q88YH6	gi|28377642	33.0/36.6	4.6/5.3	6	223	31	NAD binding	+ 3.8
Pyruvate kinase	pyk	6	Q88VY2	gi|28378548	50.0/62.9	4.7/4.99	17	1034	38	Pyruvate kinase activity	+ 2.7
Pyruvate kinase	pyk	7	Q88VY2	gi|28378548	50.0/62.9	4.8/4.99	23	1139	51	Pyruvate kinase activity	+ 4.6
Pyruvate kinase	pyk	8	Q88VY2	gi|28378548	60.0/62.9	5.75/4.99	23	1270	50	Pyruvate kinase activity	+ 3.1
Amino acid metabolism											
2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase	dapD	9	Q88V23	gi|28378858	20.0/24.5	5.8/5.02	5	193	26	Transferase activity	- 3.6
Cystathionine beta-lyase	metC2	10	Q88SB6	gi|28377181	31.5/40.8	6.15/5.52	8	351	37	Transferase activity	non detected
Cystathionine beta-lyase	metC2	11	Q88SB6	gi|28377181	31.5/40.8	6.25/5.52	2	123	6	Transferase activity	non detected
Glycine hydroxymethyltransferase	glyA	12	Q9RE02	gi|28378947	37,5/44,3	6.15/5.56	6	305	17	Pyridoxal phosphate binding	non detected
Translation											
30S ribosomal protein S6	rpsF	13	Q890K2	gi|28376981	9.2/11.4	5.55/4.97	2	141	20	rRNA binding	+ 2.2
30S ribosomal protein S6	rpsF	14	Q890K2	gi|28376981	9.2/11.4	5.35/4.97	2	101	20	rRNA binding	new induced
40S Ribosomal protein S1	rps1	15	P50889	gi|28378535	49.0/47.1	4.8/4.79	7	244	22	rRNA binding	+ 8.6
40S Ribosomal protein S1	rps1	16	P50889	gi|28378535	49.0/47.1	4.6/4.79	12	621	41	rRNA binding	+ 5.5

Table 1. *(continued)*

50S Ribosomal protein L31	rpmE2	17	Q88Z52	gi|28377400	9.0/9.1	7.15/6.82	2	70	29	rRNA binding	+ 2.2
Elongation factor G	fusA	18	Q88XY8	gi|28377830	77.5/76.9	4.8/4.81	7	370	12	GTP binding	+ 3.4
Elongation factor Tu	tuf	19	Q88VE0	gi|28378740	40.0/43.3	5.8/4.95	13	718	37	GTP binding	- 2.1
Translation initiation factor IF-1	infA	20	Q88XW4	gi|28377854	8.2/8.2	8.4/8.04	4	174	55	RNA binding	+ 2.7
Protein folding											
GroES co-chaperonin	groS	21	Q88YM6	gi|28377590	9.5/10.3	5.9/4.95	3	104	42	Cpn60 binding	+ 2
Trigger factor	tig	22	Q88VE1	gi|28378739	56.0/49.4	4.6/4.5	13	589	36	Isomerase activity	+ 4.8
Transcription											
DNA-directed RNA polymerase alpha subunit	rpoA	23	Q88XW0	gi|28377858	34.5/34.9	5.5/4.8	6	236	19	DNA binding	- 2
Other metabolic processes											
Hypothetical protein 1p_1747(Universal stress protein family like)		24		gi|28378427	13.5/18.1	5.5/4.74	4	243	34	ATP binding	- 4.3
Malolactic enzyme	mleS	22	Q88XR8	gi|28377905	56.0/59.5	4.6/4.72	5	175	14	Malic enzyme activity	+ 4.8
Short-chain dehydrogenase/oxidoreductase	yusZ	25	Q88SD1	gi|28379469	23.5/31.6	5.45/4.85	2	113	8	Oxidoreductase activity	- 2.5

[a] Spot numbers refer to those in Figure 4.
[b] Score is $-10^6\log(p)$, where p is the probability that the observed match is a random event, based on the NCBI database using the MASCOT searching program as MS/MS data.
[c] Amino acid sequence coverage for the identified protein.
[d] Fold of variation in expression in glucose *vs.* TA-grown cell: increased protein (+), decreased protein (-).

3.2.4 DISCUSSION

In this study we report that the strain VP08 of *L. plantarum* isolated from wine, displays a tannase activity. This species is the only one that shows tannase activity among wine species [25]. Moreover, analysis of growth kinetics allowed us to evaluate the physiological response of cells grown in presence of TA at enological concentrations (100 and 250 mg/L) and more (500 mg/L). Recently, we have demonstrated that, at similar experimental conditions (which mime the wine environment after alcoholic fermentation), TA produces negative effects on growth of the tannase-negative *L. hilgardii*, a wine bacterium that can co-habit with *L. plantarum* [17]. Herein, we report that growth rates of *L. plantarum*, grown in presence of TA at enological concentration (100 and 250 mg/L), are reduced, but during the stationary phase cells maintain a higher viability than glucose-grown cells. The growth test also revealed that high concentration of tannins (500 mg/L TA) resulted initially strongly inhibitory according to a previous investigation [17], with the subsequent recovering of cells viability, as a consequence of cell adaptation to tannins.

To analyze the molecular mechanisms involved in TA-mediated growth retardation and viability prolongation, we performed a proteomic analysis of starved cells grown in presence of TA or glucose as carbon source. Our data suggest that TA does not have a global effect on protein expression, but rather alters the expression of a relatively low number of proteins belonging to important cellular and metabolic pathways.

For example, here we show that 6-phosphofructokinase (pfkA), glyceraldehyde 3-phosphate dehydrogenase (gapB), and pyruvate kinase (pyk) are up-regulated in TA-grown cells during starvation. It is well known that exhaustion of an essential nutrient, and/or accumulation of a fermentation end product, limits the exponential phase of growth of microorganisms causing them to enter into the stationary phase [34, 35]. It has also been reported that during the early-stationary phase *L. plantarum* reduces the level of proteins involved in energy metabolic pathways [36]. The up-regulation of glycolytic enzymes that we report in this study suggests that TA-grown cells are able to obtain additional energy even during the starvation phase. It is possible to speculate that the tannase-positive strain VP08 of *L. plantarum* is able to gain supplementary energy by hydrolyzing TA. Indeed, hydrolysable tannins, such as TA, are composed of esters of gallic acid (gallotannins) with a sugar core which is usually glucose, and are

readily hydrolyzed by tannase into monomeric products [20]. The releasing of glucose by tannase, and the consequent upregulation of glycolytic enzymes (such as pfkA, gapB, and pyk), could correlate with the viability prolongation of TA-grown *L. plantarum* cells during the starvation phase.

Here we also show a decrease of 2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase (dapD), cystathionine beta-lyase (metC2), and glycine hydroxymethyltransferase (glyA) in TA-grown cells. DapD is a protein involved in meso-diaminopimelate/lysine biosynthetic pathway, metC2 in the degradation of cysteine and methionine [37], while glyA catalyzes the conversion of serine to glycine. The maintenance of an active metabolic state is common for LAB survival during the stationary phase [34, 35]. During starvation bacteria are capable of using alternative carbon sources such as RNA, lipids, proteins, peptides and amino acids. In particular, the catabolism of amino acids plays an important role in the survival of *L. lactis* [38] and *L. sakei* [39] during the stationary phase. In contrast with these findings our data suggest that catabolism of amino acids doesn't play a pivotal role in the survival of TA-grown *L. plantarum* cells. However, the correlation between dapD, metC2 and glyA downregulation and the prolonged viability of TA-grown cells remain to be fully clarified.

In this study, we also found that 30S ribosomal protein S6 (rpsF), 40S ribosomal protein S1 (rps1), 50S ribosomal protein L31 (rpmE2), translation initiation factor IF-1 (infA) and elongation factor G (fusA) are up-regulated in TA-grown cells. The ribosomal proteins rpsF and rpmE2 bind to 16S and 23S rRNA respectively, rps1 maintains the mRNA in a single stranded state to facilitate the 30S ribosome binding, infA regulates the 30S subunit/mRNA/initiator–tRNA complex formation, while fusA facilitates the translocation of the peptidyl-tRNA from the A-site to the P-site of the ribosome. It has been demonstrated that the bacteria growth rate and survival are directly correlated to the protein synthesizing capacity. For example, in *Lc. lactis* ribosomal proteins decreased in concentration as function of the growth rate [40]. Here we show upregulation of translation related proteins which might reflect the effect of TA on *L. plantarum* cells viability prolongation.

In this study we also report that GroES co-chaperonin (groS) and trigger factor (tig) are up-regulated in TA-grown cells. The groS is part of the chaperonin GroEL which binds, refolds, and releases denatured protein, while tig protein is a peptidylprolyl isomerase

which catalyzes protein folding, associates with nascent polypeptides on ribosomes, and cooperates with the GroEL chaperone in promoting the degradation of unstable proteins. It has been already demonstrated that both these proteins are up-regulated in heat shock response of *L. plantarum* [34]. Our data suggest that TA-grown cells are characterized by an active folding machine, which probably assure the potential to adapt to the TA stressful conditions and survive longer during starvation. Further analysis will be required on independent samples to confirm the association of the proteins we detected with the growth and survival of *L. plantarum* in wine after alcoholic fermentation.

In conclusion, this study represented the first contribution to understand the mechanisms involved in TA-mediated cell growth retardation and viability prolongation of *L. plantarum* under starvation stress (a condition that can occur frequently in a natural environment such as wine). These findings might have applications in future, especially in the field of biotechnologies, for the selection or the engineering of new malolactic starters, and for the evaluation of fitness and level of adaptation of strains under starvation stress in wine.

3.3 REFERENCES

[1] S.Q. Liu,. J. Appl. Microbiol. 92 (2002) 589-601.

[2] A. Lonvaud-Funel, Antonie Van Leeuwenhoek 76 (1999) 317-33.

[3] Y. Xu, G. Zhao, H. Pan, J.J. Li, Inst. Brew. 111 (2005) 223-228.

[4] T. Tonon, A. Lonvaud-Funel, J. Appl. Microbiol. 89 (2000) 526-531.

[5] A. Versari, G.P. Parpinello, M. Cattaneo, J. Ind. Microbiol. Biotech. 23 (1999) 447-455.

[6] G. Spano, S. Massa, Crit. Rev. Microbiol. 32 (2006) 77-86.

[7] M. Ehrnsperger, M. Gaestel, J. Buchner, Molecular Chaperones in the life Cycle of Proteins, A.L. Fink, Y. Goto (eds.), Dekker, New York, 1998.

[8] M. van de Guchte, P. Serror, C. Chervaux, T. Smokvina, S.D. Ehrlich, E. Maguin, Antonie Van Leeuwenhoek 82 (2002) 187-216.

[9] H. Alexandre, P.J. Costello, F. Remize, J. Guzzo, M. Guilloux-Benatier, Int. J. Food Microbiol. 93 (2004) 141-54.

[10] G. Spano, G. Chieppa, L. Beneduce, S. Massa, J. Appl. Microbiol. 96 (2004) 185-190.

[11] L. Beneduce, G. Spano, A. Vernile, D. Tarantino, S. Massa, J. Basic. Microbiol. 44 (2004) 10-16.

[12] E. Haslam, Practical Polyphenolics. From Structure to Molecular Recognition and Physiological Action, Cambridge University Press, Cambrige, 1998

[13] A.Scalbert, Phytochemistry 30 (1992) 3875-3883.

[14] K.T. Chung, T.Y. Wong, C.I. Wei, Y.W. Huang, Y. Lin, Crit. Rev. Food Sci. Nutr. 38 (1998) 421-464.

[15] M.R. Alberto, M.E. Farìas, M.C. Manca De Nadra, J. Food Prot. 65 (2002) 211-213.

[16] V. Renouf, O. Claisse, C. Miot-Sertier, A. Lonvaud-Funel, Food Microbiol. 23 (2006) 136-145.

[17] A. Bossi, S. Rinalducci, L. Zolla, P. Antonioli, P.G. Righetti, G. Zapparoli, J. Appl. Microbiol. 102 (2007) 787-795.

[18] L. Mingshu, Y. Kai, H. Qiang, J. Dongying, J. Basic Microbiol. 1 (2006) 68-84.

[19] C. Aguilar, R. Rodrìguez, G. Gutiérrez-Sanchez, C. Augur, E. Favela-Torres, L.A. Prado-Barragan, A. Ramírez-Coronel, J.C. Contreras-Esquivel, Appl. Microbiol. Biotechnol. 76 (2007) 47-59.

[20] A.M. Deschamps, G. Otuk, J.M. Lebeault, J. Ferment. Technol. 61 (1983) 55-59.

[21] K.C. Mondal, R. Banerjee, B.R. Pati, Biotechnol. Lett. 0 (2000) 767-769.

[22] R. Osawa, K. Kuroiso, S. Goto, A. Shimizu, Appl. Environ. Microbiol. 66 (2000) 3093-3097.

[23] Y. Nishitani, E. Sasaki, T. Fujisawa, R. Osawa, Syst. Appl. Microbiol. 27 (2004) 109-117.

[24] A. Sabu, C. Augur, C. Swati, A. Pandey, Process Biochem. 41 (2006) 575-580.

[25] I. Vaquero, A. Marcobal, R. Muñoz, Int. J. Food Microbiol. 96 (2004) 199-204.

[26] E. Sasaki, T. Shimada, R. Osawa, Y. Nishitani, S. Spring, E. Lang, Syst Appl Microbiol. 28 (2005) 358-65.

[27] S.P. Chowdhury, S. Khanna, S.C. Verma, A.K. Tripathi, J. Appl. Microbiol. 97 (2004) 1210-1219.

[28] L. Ayed, M. Hamdi, Biotechnol. Lett. 24 (2002) 1763-1765.

[29] H. Rodrìguez, B. De Las Rivas, C. Gòmez-Cordovéz, R. Muñoz, Int. J. Food Microbiol. 121 (2008) 92-98.

[30] P.K. Lekha, B. Lonsane, Adv. Appl. Microbiol. 44 (1997) 215-260.

[31] A.L. Waterhouse, Ann N Y Acad Sci. 957 (2002) 21-36.

[32] Y. Nishitani, R. Osawa, J. Microb. Methods 54 (2003) 281-284.

[33] A. Shevchenko, M. Wilm, O. Vorm, M. Mann, Anal. Chem. 68 (1996) 850–858.

[34] F. Al-Shahrour, P. Minguez, J. Tárraga, D. Montaner, E. Alloza, J.M. Vaquerizas, L. Conde, C. Blaschke, J. Vera, J. Dopazo, Nucleic Acids Research 34 (2006) 850-858.

[34] M. De Angelis, M. Gobbetti, Proteomics, 4 (2004) 106-122.

[35] B. Ganesan, M. Stuart, B.C. Weimer, Appl. Environ. Microbiol. 73 (2007) 2498-2512.

[36] D.P. Cohen, J. Renes, F.G. Bouwman, E.G. Zoetendal, E. Mariman, W.M. de Vos, E.E. Vaughan, Proteomics 6 (2006) 6485-93.

[37] S. Irmelr, S. Raboud, B. Beisert, D. Rauhut, H. Berthoud, Appl. Environ. Microbiol. 74 (2008) 99-106.

[38] M.R. Stuart, L.S. Chou, B.C. Weimer, Appl. Environ.Microbiol. 65 (1999) 665–673.

[39] M.C. Champomier Vergès, M. Zúniga, F. Morel-Deville, G. Pèrez-Martinez, M. Zagorec, S.D. Ehrlich, FEMS Microbiol. Lett. 180 (1999) 297–304.

[40] J. Palmfeldt, F. Levander, B. Hahn-Hägerdal, P. James, Proteomics 4 (2004) 3881-3898.

CHAPTER 4

PROTEOMIC ANALYSIS OF PLANT-PATHOGEN INTERACTION

4.1 INTRODUCTION

Both aerial and subterranean plant organs are constantly exposed to intimate contacts with a plethora of viruses and microrganisms, as diverse as bacteria, oomycetes, fungi, and eukaryotic protozoans. The outcome of interactions between plants and microbes can be neutral, detrimental or even beneficial for the photoautotrophic organisms. Disadvantageous encounters typically result in disease, which in extreme cases can lead to full collapse of plant tissues. Plant pathogens such as bacteria, fungi and viruses have the potential to cause (and have historically caused) devastating crop losses worldwide. Plant–pathogen interactions have been studied extensively over the years from both the plant and pathogen viewpoints. Looking at the defence mechanisms in plants, the recognition and signalling events that occur in plant cells in response to microorganisms need to be extremely rapid, reliable and specific, and are part of the strategy evolved by plants to survive attacks. The recognition of pathogen-associated molecular patterns (PAMPs) [1] or pathogen effectors lead to the activation of the plant basal defence [2]. Therefore, successful pathogens must evolve mechanisms to avoid recognition or to interfere with basal defence to colonize the host and develop disease. Superimposed on the basal defence, some plant varieties express resistance proteins that guard against this interference and trigger a specific, genetically defined response. Irrespective of the type of pathogen recognition, plants can activate a wide array of antimicrobial responses. The earlier events are changes in ion fluxes, protein phosphorylation, nitric oxide generation, a burst of reactive oxygen species and lipid hyperperoxidation, followed by signal transduction. A battery of defence genes are then activated, which lead to biosynthesis of pathogenesis-related proteins (PR), antimicrobial phenolic compounds and to plant cell wall thickening and papilla deposition [3].

With regard to plant pathogens, the capacity to protect themselves from the oxidative stress activated by the plant is of extreme importance. Therefore, pathogens produce several genes, such as catalases and superoxide dismutase, which are responsible for the inactivation of H_2O_2 and O_2^-. The importance of secretion pathways for pathogenicity has also been well established. The pathogens expresse and secrete degrading enzymes such as pectin esterases, polygalacturonases, xylanases, pectato lyases and cellulases, toxins, and effector proteins that are predicted to collaborate in the suppression of basal resistance through the modification of specific host proteins [4].

These aspects of plant-pathogen interaction have been mainly investigated in well-known plant model systems, such as *Arabidopsis thaliana*, tobacco, rice and a few others, and the derived information can only partially be extended to crop species. In particular, signals and mechanisms operating in grapevine against pathogens are still largely unknown.

4.2 PROTEOMIC ANALYSIS OF GRAPEVINE LEAVES AFTER *PLASMOPARA VITICOLA* INFECTION

4.2.1 *VITIS VINIFERA-PLASMOPARA VITICOLA* CHALLENGE

Grapevine (*Vitis vinifera*) is one of the world's more cultivated fruit crop. Unfortunately, all cultivars currently used for grape or wine production are susceptible to several pathogens, such as the downy mildew pathogen, *Plasmopara viticola*. Resistant sources are available, but breeding efforts are still in progress [5] to produce new resistant cultivars with good organolectic properties.

Downy mildew is a devastating disease that provokes significant economic losses, as well as environmental damage due to the repetitive applications of fungicides. Primary infection begins with over-wintering oospores, in infected leaves, on the ground and possibly in diseased shoots, which germinate in the spring and produce sporangia. Successively, the sporangia release zoospores, which produce a germ tube able to enter through stomata. Once inside the plant, the mycelium grows and spreads through tissues, causing visible lesions in about 5-12 days, which appear as yellow to grayish-

yellow spots on the upper surface (so-called “oil spots”). Following a variable incubation period and in suitable environmental conditions, the pathogen can produce secondary inoculum in form of sporangiophores growing out through stomata on the lower leaf surface which are spread by rain causing infections in distant sites. These appear as white to grayish cotton-like fungal growth on the lower surface. In late season infections, small, angular, yellow to red-brown spots develop on the upper surface and lesions commonly form along the veins. Infected berries turn brown and soft, shatter easily, and often are covered with pathogen structures. Young shoots usually are stunted and become thickened and distorted. Without any control downy mildew may lead to crop losses of 50-75%. After a variable number of secondary infections, the cycle ends with the sexual production of over-wintering oospores.

Molecular aspects of the infection process are still poorly understood, but recent findings suggest that signals are exchanged between the pathogen and host to establish and maintain compatibility and possibly to suppress defense responses [6]. Secreted virulence factors may be involved in this process [7] and four such gene products have recently been identified in other oomycetes [8].

The plant's response to infection has been characterized predominantly through the study of incompatibility in the resistant species *Vitis riparia, V. rupestris and Muscadinia rotundifolia*. These previous data with downy mildew-resistant in vitro plants had shown that infection with *P. viticola* was associated with an expression of distinct reactions in a chronological order, such as increased production of reactive oxygen species, hypersensitive response, increased peroxidase activity, and accumulation of phenolic compounds [9]. Further studies have elucidated the accumulation of diverse group of phenolic compounds such as lignins, coumarins and stilbenes. Mainly, the deposition of lignin [4] and other phenolic compounds, accompanied by high peroxidase activity, and the production of stilbene phytoalexins represent important components of a defence response in grapevine. Preformed isoflavonoids (flavonoid derivatives), and induced flavonoids, may also play a role in protecting the young berries or leaves from various phytopathogens, such as *P. viticola* [10]. In addition, plants prevent pathogen infections by the synthesis of a number of pathogenesis-related (PR) proteins [11].

Molecular data obtained by the direct investigation of compatible interactions in cultivated grapevine genotypes is scarce. The study of *V. vinifera* susceptibility may be

helpful to understand the factors potentially involved in *P. viticola*-induced disease. As theoretical scenario these findings may have consequences in terms of the generation of transgenic or mutant plants whose endogenous levels of crucial proteins are affected, thereby conferring increased tolerance to this oomycete phytopathogen.

Plant, pathogens and their interactions have been studied using classical biochemical, genetic, molecular biological and plant pathology approaches. New approaches (such as transcriptomics and proteomics) are essential to provide global information on the various cellular genomic and proteomic networks, which could be finally combined in system biology approaches [12]. In particular, proteomics provides a global approach to explore changes in abundance of specific components of the pathosystem proteome, and hence to identify specific proteins and processes likely to be central to the outcome of infection. Nevertheless, to date, few proteomic studies of oomycete pathogens and their interaction with host plants have been performed. Most have focused on *Phytophthora* species such as *P. infestans*, a devastating pathogen of solanaceous plants [13, 14], *P. nicotianae* which has a wide host range [15] and *P. palmivora*, a serious pathogen of tropical crops including cocoa.

Up to date, proteomic data about grapevine-oomycete interactions are almost completely missing. In this study, a comparative proteomic approach to characterize differentially expressed proteins of susceptible *V. vinifera* infected by *P. viticola* has been developed for the first time. The present work provides fundamental findings regarding plant proteins specifically involved in a compatible interaction.

4.2.2 MATERIALS AND METHODS

4.2.2.1 Plant Material, Inoculum And Pathogen Infection

Grapevine plants (cv Pinot Noir) were grown in a greenhouse at 24°C with 50% relative humidity. The *P. viticola* inoculum was collected from sporulated field leaves and used for the artificial inoculations of surface-sterilized leaves. The inoculum was stored as sporangia at -20°C. Leaves of Pinot Noir from the greenhouse were posed in Petri plates, and infection was initiated by deposing droplets of a suspension of 30000

sporangia per ml in cold pure water on the lower surface of the leaves. Samples were incubated and collected at 24, 48 and 96 hours post-infection (hpi). As control, leaves were treated with water and collected at the same time points.

4.2.2.2 Protein Extraction

Leaves of Pinot Noir infected by *P. viticola* or treated with water (for 24, 48 and 96 hours), were homogenized to a very fine powder in a liquid nitrogen-precooled mortar by using a pestle. Total protein extraction was performed by adding 5 ml (per gram of tissue) of a solution containing 7 M urea, 2 M thiourea, 80 mM citric acid, 1% C7Bz0 and 1x protease inhibitor cocktail tablets. Samples were maintained in ice for one hour (vortexing each 10 minutes) and then centrifuged 30 mins at 40000 x g at 4°C. The supernatants were collected and the centrifugation was repeated as above. Proteins were precipitated in acetone:methanol 8:1 at -20°C over-night and pelleted by centrifugation at 17500 x g for 10 mins. The pellets were washed three times with acetone:methanol 50:50 and the proteins were solubilised with a solution containing 7 M urea, 2 M thiourea, 3% CHAPS, 40 mM Tris, 1% pH 3-10 Ampholyte. After 1 h of incubation, the samples were centrifugedated 30 min at 17500 x g to remove debris and the nucleic acids complexed with ampholytes. The protein concentration was evaluated by the Bradford assay. The samples were incubated with 5 mM tributyl phosphine and 20 mM acrylamide for 60 minutes at room temperature to reduce protein disulphide bonds and alkylate the cysteine thiolic groups. The reaction was blocked by the addition of 10 mM DTT and the samples were collected and stored at -80°C.

4.2.2.3 2-D Gel Electrophoresis

Two-dimensional gel electrophoresis was performed as described in Section 2.3.2.5.

4.2.2.4 Image Analysis

Gel scanning and image analysis of six experimental groups (24, 48 and 96 h *P. viticola* or water treated samples) were performed as decribed in Section 2.3.2.6. A Student's t-test was performed in order to compare the differences beetwen control and 24, 48 and 96 h pathogen treated samples. Sets of proteins that showed a statistically significant difference with a confidence level of 0.05 were identified.

4.2.2.5 In-Gel Digestion

Spots showing a statistically significant differential expression were carefully manually cut out from 2-D Sypro Ruby stained gels and in gel digestion was performed as described in Section 2.3.2.7.

4.2.2.6 Peptide Sequencing By Nano-RP-HPLC-ESI-MS/MS

Peptide mixtures mass spectrometric analysis and protein identification were performed as described in Section 3.2.2.8.

4.2.3 RESULTS

In this study 2-DE analysis was performed to evaluate the effects of the pathogen infection on the leaf proteome of *V. vinifera* cultivar Pinot Noir. Five independent gels were obtained for each sample (24, 48 and 96 h pathogen or water treated samples), (Fig 1 and 2) and more than 500 spots were resolved in the 3–10 pH range. In particular, 505±23 spots have been detected in the control group and 537±25 spots in the infected one. The representative images of 2-D maps are shown in Fig 3A, B and C, together with enlarged images of some regulated spots ($p<0.05$). Based on PDQuest analysis, 100 spots (marked by red crosses in Fig. 4) were observed to be modulated (at least 2 fold) in Pinot Noir following pathogen challenge. In particular, 81 spots were increased in intensity, whereas 19 decreased. Among them, 53 spots were regulated at 24 hpi (42 up-regulated and 11 down-regulated), while 14 at 48 hpi (9 up-regulated and 5 up-regulated) and 47 at 96 hpi (42 up-regulated and 5 down-regulated). All these spots were excised from the gels and subjected to RP-HPLC-ESI-MS/MS analysis for protein identification. A total of 103 proteins were identified (because of the presence of more than one protein in some spots), corresponding to 72 unique polypeptides. In Table 1, their identities are shown, together with the standard spot number (SSP), the protein databases ID, and the fold variation after the different treatments (24h, 48h, and 96h). Not all the differentially expressed spots were identified because of their relative low concentrations.

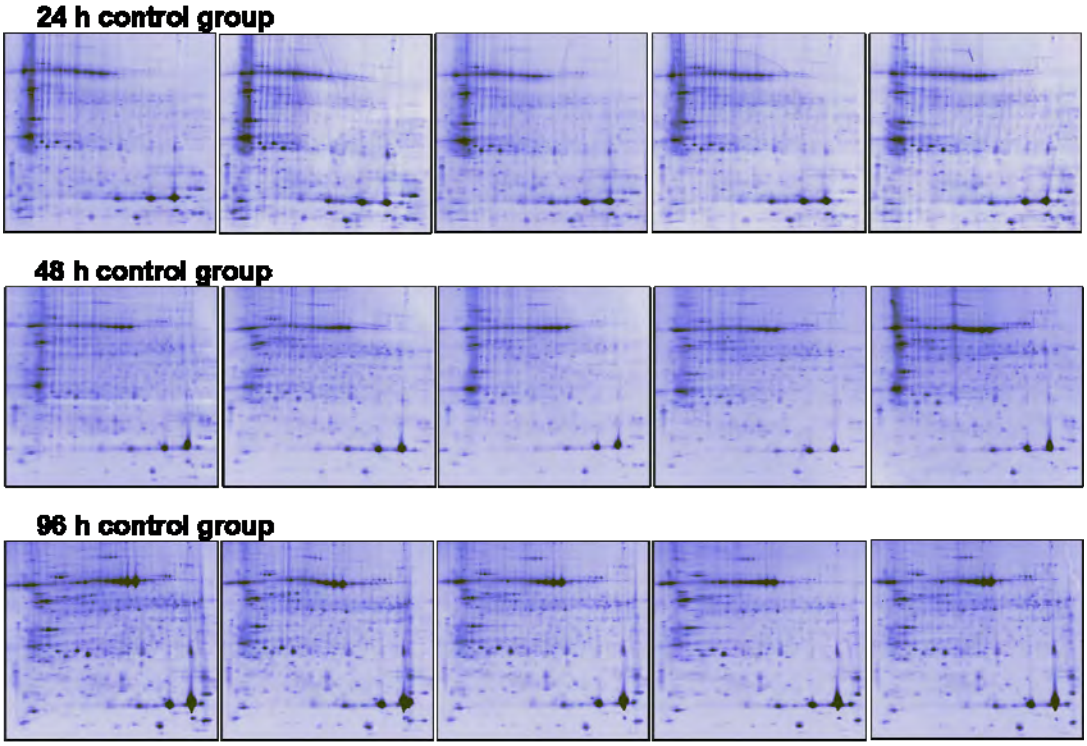

Fig. 1. Water treated samples 2-D maps. Five independent gels were obtained for each treatment: 24, 48 and 96 h.

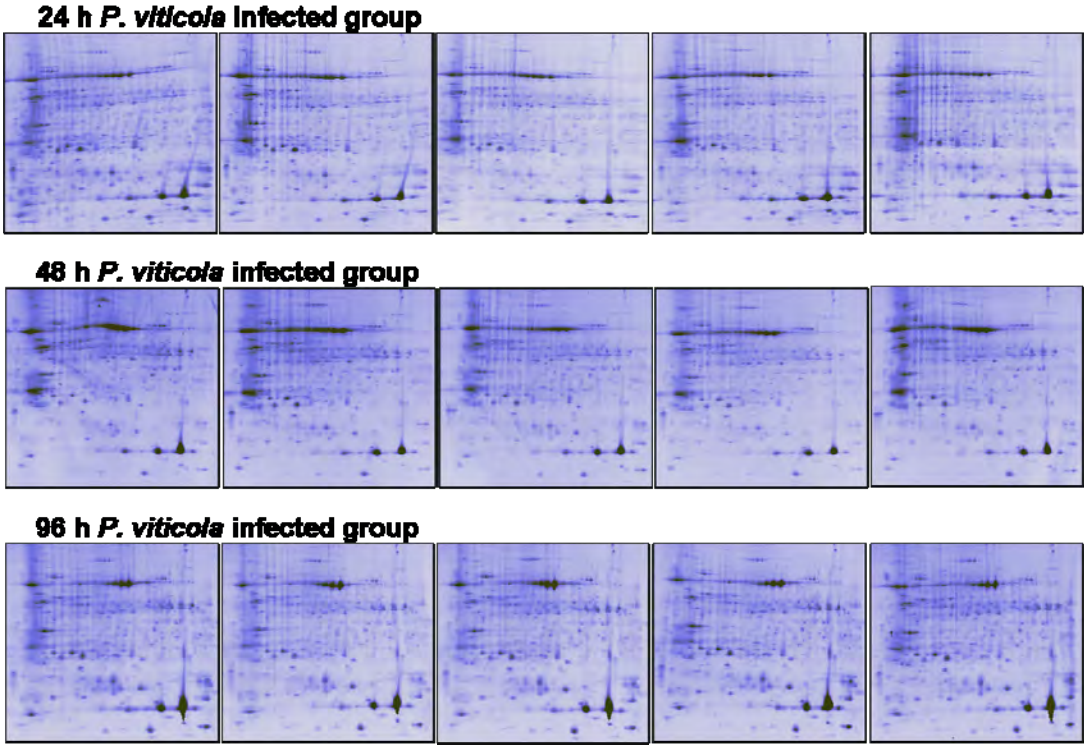

Fig. 2. Pathogen infected samples 2-D maps. Five independent gels were obtained for each treatment: 24, 48 and 96 h.

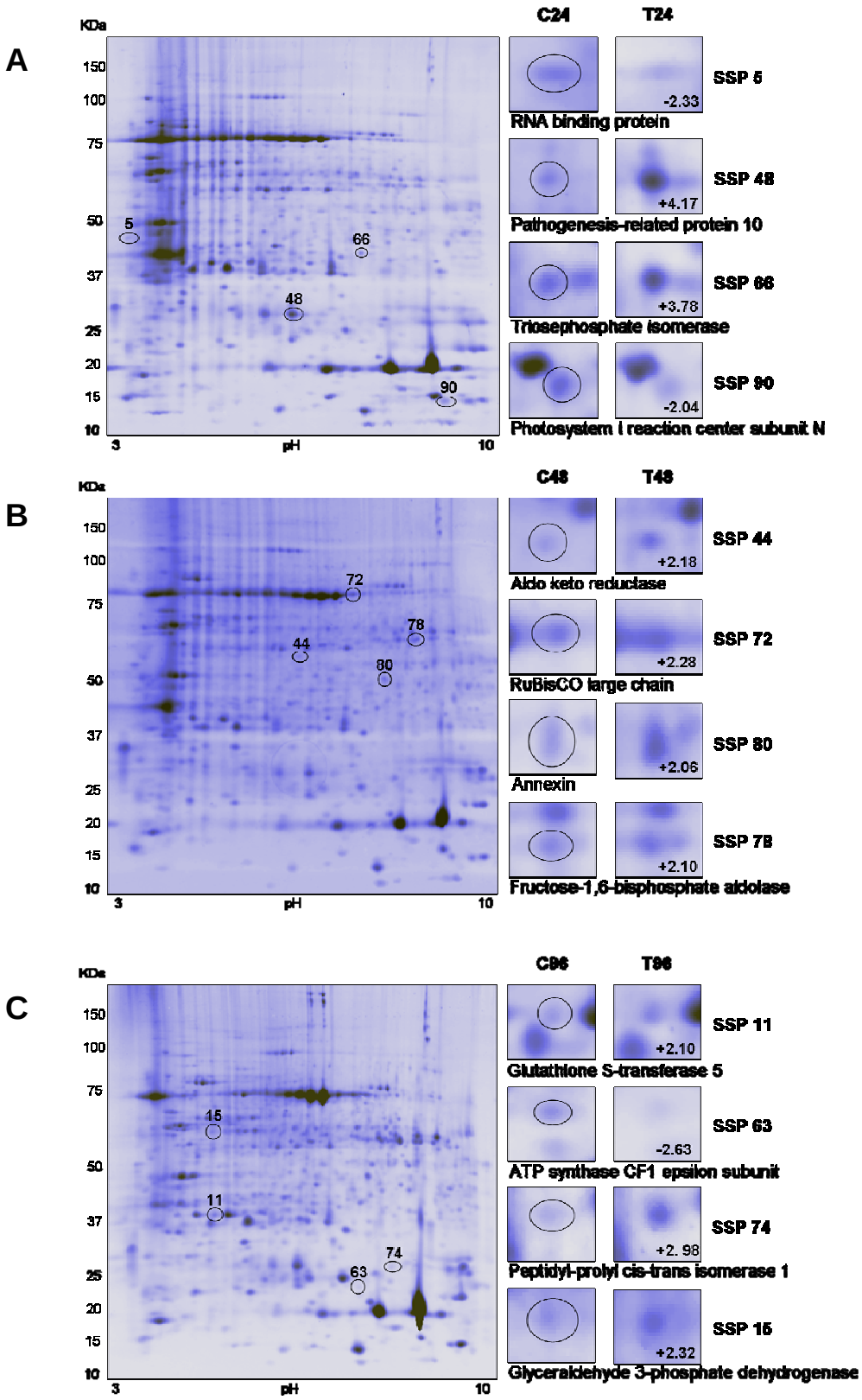

Fig. 3. High-resolution two-dimensional gel electrophoresis. Example of Sypro Ruby stained 2-D gels of proteins extracted from Pinot Noir leaves after (A) 24 h, (B) 48 h and (C) 96 h after pathogen inoculation, and PDQuest (version 7.3) output showing representative differentially expressed spots (p <0.05) between control and infected samples. For each spot, an enlarged region of the respective 2-DE map is shown which contains the referred spot (highlighted by an ellipse) and the corresponding fold of change. Each spot was identified by MS/MS analysis.

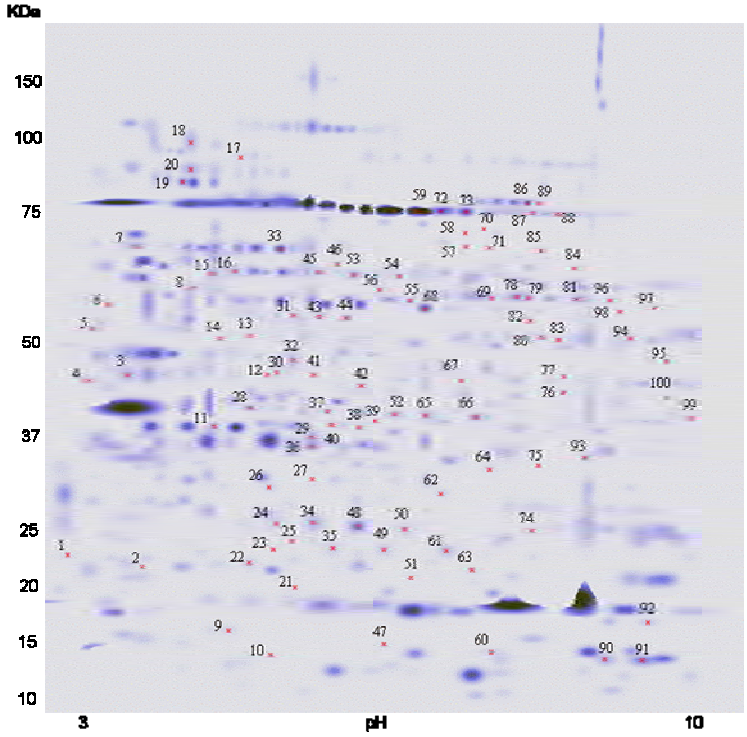

Fig. 4. The standard map is represented with all the 100 spots (marked by red crosses) found to be differentially expressed after infection with *P. viticola.*

Table 1. Identified differentially expressed proteins by nano-RP-HPLC-ESI-MS/MS analysis

Protein name	Gene name	SSP	SwissProt / TrEMBL #	NCBI #	Fold of variation (C *vs.* 24h)	Fold of variation (C *vs.* 48h)	Fold of variation (C *vs.* 96h)	Molecular function GO term	Cellular component GO term
Oxidative stress-related proteins									
Aldo keto reductase	GSVIVT00021382001	44	A7PNT4	gi|157346009	2,58	2,18		oxidoreductase activity	Cytosol
Catalase	GSVIVT00002880001	87	A7QPB7	gi|157358559	2,17			catalase activity	Peroxisome
Catalase	GSVIVT00002880001	57	A7QPB7	gi|157358559	new induced			catalase activity	Peroxisome
Catalase	GCat	86	Q8S568	gi|19070130	2,11			catalase activity	Peroxisome
Catalase	GCat	89	Q8S568	gi|19070130	2,44	2,14		catalase activity	Peroxisome
Copper chaperone	GSVIVT00034927001	9	A7QBE3	gi|157353972			2,87	metal ion binding activity	Chloroplast
Ferredoxin--NADP reductase	VITISV_042164	77	A5BG49	gi|147791392			2,29	oxidoreductase activity	Chloroplast
Flavodoxin 1	VITISV_016391	40	A5AS18	gi|147788048		2,13		oxidoreductase activity	Chloroplast
FMN-dependent dehydrogenase	VITISV_021217	100	A5B9Z0	gi|147770035	4,65			oxidoreductase activity	Cytosol
Glutathione peroxidase [*Picea sitchensis*]		62	A9NQB1	gi|116783172			2,06	peroxidase activity	Cytosol
Glutathione S-transferase	GSVIVT00027957001	29	A7Q089	gi|157350036	2,02			transferase activity	Cytosol
Glutathione S-transferase 5	VITISV_036241	11	A5ARN9	gi|147781552			2,10	transferase activity	Cytosol
PII nitrogen sensing protein GLB I	GSVIVT00028171001	22	A7Q0T3	gi|157350230		-2,13		enzime regulator activity	Choloroplast
Porin 3	GSVIVT00028995001	67	A7Q2S2	gi|157350923			2,86	trasporter activity	Mithocondrion
Prohibitin 1	VITISV_018854	95	A5BVI0	gi|147791337			8,56	protein binding	Mitochondrion
Protein disulfide isomerase	VITISV_038267	16	A5C7J5	gi|147821099			3,66	isomerase activity	Chloroplast
Secretory peroxidase	GSVIVT00023967001	97	A7NY33	gi|157336988	8,04		new induced	oxidoreductase activity	Secreted
Sulfite reductase, alpha subunit	GSVIVT00014406001	83	A7NSU2	gi|157335145	2,26			oxidoreductase activity	Chloroplast
Thiolase	VITISV_023191	71	A5BXL8	gi|147866537			2,20	transferase activity	Mitochondrion , peroxisome
Thiolase	VITISV_023191	88	A5BXL8	gi|147866537	2,39			transferase activity	Mitochondrion , peroxisome
Vacuolar ATPase subunit A	GSVIVT00029622001	18	A7Q4B5	gi|157351469			2,69	ATPase activity	Vacuole

Table 1. *(continued)*

Defense-related proteins									
Basic endochitinase	CHIT1B	99	P51613	gi|1705813			2,19	chitinase activity	Membrane
Dabb, Stress responsive domain	GSVIVT00011995001	10	A7R8M0	gi|157333599	-2,50			unknown	Unknown
Pathogenesis-related protein 10	PR10.2 (VITISV_017150)	23	Q9FS43	gi|11182124	-2			ribonuclease activity	Cytosol
Pathogenesis-related protein 10	PR10.1	25	Q9FS42	gi|11182126			8,00	ribonuclease activity	Cytosol
Pathogenesis-related protein 10	PR10.2 (VITISV_017150)	34	Q9FS43	gi|11182124	4,47		2,16	ribonuclease activity	Cytosol
Pathogenesis-related protein 10	PR10.2 (VITISV_017150)	35	Q9FS43	gi|11182124			2,78	ribonuclease activity	Cytosol
Pathogenesis-related protein 10	PR10.2 (VITISV_017150)	50	Q9FS43	gi|11182124			2,33	ribonuclease activity	Cytosol
Pathogenesis-related protein 10	PR10	61	Q2I305	gi|86156032			3,44	ribonuclease activity	Cytosol
Pathogenesis-related protein 10.1	PR10.1	49	Q20BD3	gi|89887945			4,20	ribonuclease activity	Cytosol
Pathogenesis-related protein 10	PR10.2 (VITISV_017150)	48	Q9FS43	gi|11182124	4,17		1,96	chitinase activity	Cytosol
Pathogenesis-related protein Bet v 1	GSVIVT00033082001 GSVIVT00033084001	1	A7Q8J9	gi|157352968	-2,70			ribonuclease activity	Cytosol
Similar to Macrophage migration inhibitory factor	GSVIVT00037405001	60	A7QFF7	gi|157355400			2,07	isomerase activity	Unknown
Photosynthesis									
ATP synthase CF1 alpha subunit	atpA	19	Q0ZJ35	gi|91983977			-2,27	ATPase activity	Chloroplast
ATP synthase CF1 epsilon subunit	atpE	22	Q0ZJ14	gi|91983998		-2,13		ATPase activity	Chloroplast
ATP synthase CF1 epsilon subunit	atpE	63	Q0ZJ14	gi|91983998		-2,27	-2,63	ATPase activity	Chloroplast
Photosystem I reaction center subunit N	PsaN	90	Q6XGX6	gi|37927472	-2,04			protein binding	Chloroplast
Photosystem I reaction center subunit N	PsaN	91	Q6XGX6	gi|37927472			-2,04	protein binding	Chloroplast
Photosystem II stability/assembly factor HCF136	GSVIVT00030205001	8	A7P2W0	gi|157338667	new induced		2,22	hydrolase activity	Chloroplast
PsbP	VITISV_028610	26	A5B1G1	gi|147787657	2,24			ion binding	Chloroplast
PsbP	VITISV_001116	36	A5B1D3	gi|147787750	-2,17			ion binding	Chloroplast
Calvin cycle and glycolysis									
Carbonic anhydrase	GSVIVT00020732001	39	A7PM94	gi|157345467		2,38		lyase activity	Chloroplast
Carbonic anhydrase	GSVIVT00020732001	37	A7PM94	gi|157345467			2,22	lyase activity	Chloroplast
Carbonic anhydrase	GSVIVT00020732001	29	A7PM94	gi|157345467	2,02			lyase activity	Chloroplast
Carbonic anhydrase	GSVIVT00020732001	38	A7PM94	gi|157345467	2,23			lyase activity	Chloroplast
Carbonic anhydrase	GSVIVT00020732001	42	A7PM94	gi|157345467	2,19			lyase activity	Chloroplast
Carbonic anhydrase	GSVIVT00020732001	65	A7PM94	gi|157345467	2,72			lyase activity	Chloroplast
Carbonic anhydrase	GSVIVT00020732001	40	A7PM94	gi|157345467		2,13		lyase activity	Chloroplast
Fructose-1,6-bisphosphate aldolase	VITISV_035583	96	A5B118	gi|147781269			4,09	lyase activity	Chloroplast
Fructose-1,6-bisphosphate aldolase	VITISV_020600	68	A5BDH7	gi|147784332	2,09			lyase activity	Chloroplast
Fructose-1,6-bisphosphate aldolase	VITISV_035583	78	A5B118	gi|147781269	2,93	2,10		lyase activity	Chloroplast
Glyceraldehyde 3-phosphate dehydrogenase	GSVIVT00009717001	79	A7QYF8	gi|157327869			2,87	oxidoreductase activity, NAD binding	Cytosol
Glyceraldehyde 3-phosphate dehydrogenase	VITISV_021726	15	A5BCS8	gi|147858738			2,32	oxidoreductase activity, NAD binding	Cytosol
Glyceraldehyde 3-phosphate dehydrogenase	VITISV_021726	45	A5BCS8	gi|147858738	3,26			oxidoreductase activity	Cytosol
Glyceraldehyde 3-phosphate dehydrogenase	VITISV_021726	53	A5BCS8	gi|147858738	5,26			oxidoreductase activity	Cytosol
Glyceraldehyde 3-phosphate dehydrogenase	VITISV_021726	54	A5BCS8	gi|147858738	5,55			oxidoreductase activity	Cytosol

Table 1. *(continued)*

Glyceraldehyde 3-phosphate dehydrogenase	GSVIVT00009717001	78	A7QYF8	gi|157327869	2,10			oxidoreductase activity	Cytosol
RuBisCO large chain	rbcL	77	Q6ZYB1	gi|17232956			2,29	monooxygenase activity	Chloroplast
RuBisCO large chain	rbcL	59	Q6ZYB1	gi|46275228	3,07			monooxygenase activity	Chloroplast
RuBisCO large chain	rbcL	72	Q6ZYB1	gi|46275228	2,42	2,28		monooxygenase activity	Chloroplast
RuBisCO large chain	rbcL	73	Q6ZYB1	gi|46275228	2,55			monooxygenase activity	Chloroplast
RuBisCO large chain	rbcL	98	Q6ZYB1	gi|46275228			1,98	monooxygenase activity	Chloroplast
RuBisCO large chain	rbcL	17	P00875	gi|11497536			3,39	monooxygenase activity	Chloroplast
RuBisCO large chain	rbcL	27	P00876	gi|11497536			-2,50	monooxygenase activity	Chloroplast
RuBisCO small chain	VITISV_022329	21	A5C718	gi|147807983	-2,17			monooxygenase activity	Chloroplast
RuBisCO small chain	VITISV_022329	93	A5C718	gi|147807983	-5,88			monooxygenase activity	Chloroplast
Triose-phosphate isomerase	VITISV_041523	37	A5BV65	gi|147784332			2,22	isomerase activity	Chloroplast
Triose-phosphate isomerase	VITISV_041523	27	A5BV65	gi|147784332			-2,50	isomerase activity	Chloroplast
Triose-phosphate isomerase	VITISV_041523	52	A5BV65	gi|147784332	2,12			isomerase activity	Chloroplast
Triose-phosphate isomerase	VITISV_041523	66	A5BV65	gi|147784332	3,78			isomerase activity	Chloroplast
Citric acid (Krebs) cycle									
Citrate synthase	GSVIVT00003727001	88	A7QR94	gi|157359248	2,39			transferase activity	Mitochondrion
Fumarase class II	GSVIVT00028048001	58	A7Q0H2	gi|157350119			2,74	lyase activity	Mitochondrion
Malate dehydrogenase	VITISV_033330	94	A5BEJ8	gi|147774268			-2,63	oxidoreductase activity	Mitochondrion
RNA Transcription and Translation									
Elongation factor Tu	VITISV_014296	46	A5BV56	gi|147784261			3,34	translation factor activity, nucleic acid binding	Cytosol
Ribosomal protein L1	GSVIVT00028715001	94	A7Q228	gi|157350678			-2,63	structural costituent of ribosome	Ribosome
RNA binding protein	GSVIVT00015937001	5	A7PCN3	gi|157342097	-2,33			nucleotide binding	Cytosol
Transcription factor homologous to NACalpha-BTF3	GSVIVT00023323001	4	A7PTE4	gi|157347628		2,65		protein binding	Nucleus
Pyridoxamine 5'-phosphate oxidase	GSVIVT00025470001	32	A7PW46	gi|157348585	-2,08			cofactor binding	Cytosol
Translation initiation factor IF-3	GSVIVT00017183001	76	A7PFP1	gi|157343157	2,22		2,12	translation factor activity	Cytosol
Protein folding and metabolism									
Chaperonin 21	cpn21	28	Q6B4V4	gi|50660327	2,1			ATP binding	Chloroplast
Chaperonin GroEL-related	GSVIVT00024837001	98	A7P0B1	gi|157337766			1,98	protein binding	Chloroplast
Chaperonin GroEL-related	GSVIVT00024837001	20	A7P0B1	gi|157337766			2,02	protein binding	Chloroplast
Peptidyl-prolyl cis-trans isomerase 1	VITISV_030108	74	A5AKD8	gi|147787082	2,05		2,98	isomerase activity	Chloroplast
Peptidyl-prolyl cis-trans isomerase 1	GSVIVT00036804001	92	A7QE01	gi|157354889			2,48	isomerase activity	Chloroplast
Proteasome alpha type 7	GSVIVT00036109001	76	A7P8A6	gi|157340566	2,22		2,12	peptidase activity	Proteasome core complex
Metabolic processes (pentose phosphate shunt, mevalonate pathway, etc…)									
6-phosphogluconate dehydrogenase decarboxylating	VITISV_032572	70	A5B7A4	gi|157350923		3,04	3,49	oxidoreductase activity	Cytosol
Acetyl-CoA acetyl transferase	VITISV_023191	85	A5BXL8	gi|147866537	4,20			transferase activity	cytosol, mitochondrion peroxisomes
Cysteine synthase	VITISV_033255	8	A5AFH5	gi|147819267	new induced		2,22	lyase activity, transferase activity	cytosol, chloroplast
Inorganic pyro-phosphatase	GSVIVT00024902001	28	A7PUS5	gi|157348112	2,1			hydrolase activity, ion binding	Vacuole
Nucleoside-diphosphate-sugar epimerase	GSVIVT00036212001	43	A7P8K3	gi|157340663	2,03			catalytic activity	Nucleus
Nucleoside-diphosphate-sugar epimerase	GSVIVT00024587001	31	A7NZN9	gi|157337544			2,04	catalytic activity	Nucleus
Tropine dehydrogenase	GSVIVT00018424001	76	A7PIP6	gi|157344215	2,22		2,12	oxidoreductase activity	Cytosol
UTP--glucose-1-phosphate uridylyltransferase	GSVIVT00034877001	73	A7QBA6	gi|157353934	2,55			transferase activity	Cytosol

Table1. *(continued)*

Others

Annexin	VITISV_020506	80	A5BTZ8	gi|147861246		2,06		calcium ion binding	Membrane
C2 domain-containing protein	VITISV_006805	6	A5C8Y4	gi|147836228			2,17	phospholipid binding	Membrane
Cold Shock Protein 2	VITISV_040408	64	A5BQ96	gi|147856673	2,72		2,13	DNA binding	Nucleus
Hexapeptide transferase family protein	VITISV_014473	30	A5C3G7	gi|147856360		2,14		transfarase activity	Cytosol
PAP fibrillin	GSVIVT00026573001	75	A7PYT8	gi|157349531			2,14	structural molecule activity	Plastid
PAP fibrillin	VITISV_040863	3	A5B0L5	gi|147856329		-3,22		structural molecule activity	Plastid
SAR1/GTP-binding secretory factor	GSVIVT00033116001	64	A7Q8N0	gi|157352999	2,72		2,13	GTP binding activity	Membrane
Thylakoid lumenal 17.4 kDa protein	GSVIVT00037228001	2	A7QF07	gi|157355248	-2,78			unknown	Chloroplast
WD40 repeat protein	VITISV_030908	82	A5AL04	gi|147853445	2,22			protein binding	Cytosol

Functional annotation of the proteins was carried out by categorizing them into different groups based on GO terms. Figures 4 shows pie-chart distributions of the identified proteins, catalogued according to the biological process in which they are involved. Taken together (top-left chart), out of the 72 unique identified proteins identified, 28% were involved in Calvin cycle and glycolysis, 19% in oxidative stress response, 12% in defense response, and 9% in photosynthesis and 6% in protein folding. This distribution is also closely mirrored in samples collected at 24 hpi (top-right chart). Conversely, at 48 hpi the number of identified modulated proteins is much lower (16) and the distribution in functional categories also differs from the first collection time (middle-right chart): no differentially expressed proteins involved in defense response and protein folding were detected, whereas a relatively higher number (18%) of photosynthesis-related proteins was found. Finally, at 96 hpi the number of identified modulated proteins rises again to 51 (bottom-right chart), and defence-related proteins are again consistently detected (18% of identified proteins) together with 19% involved in oxidative stress response, 19% in the Calvin cycle and glycolysis functional category, and 10% involved in protein folding.

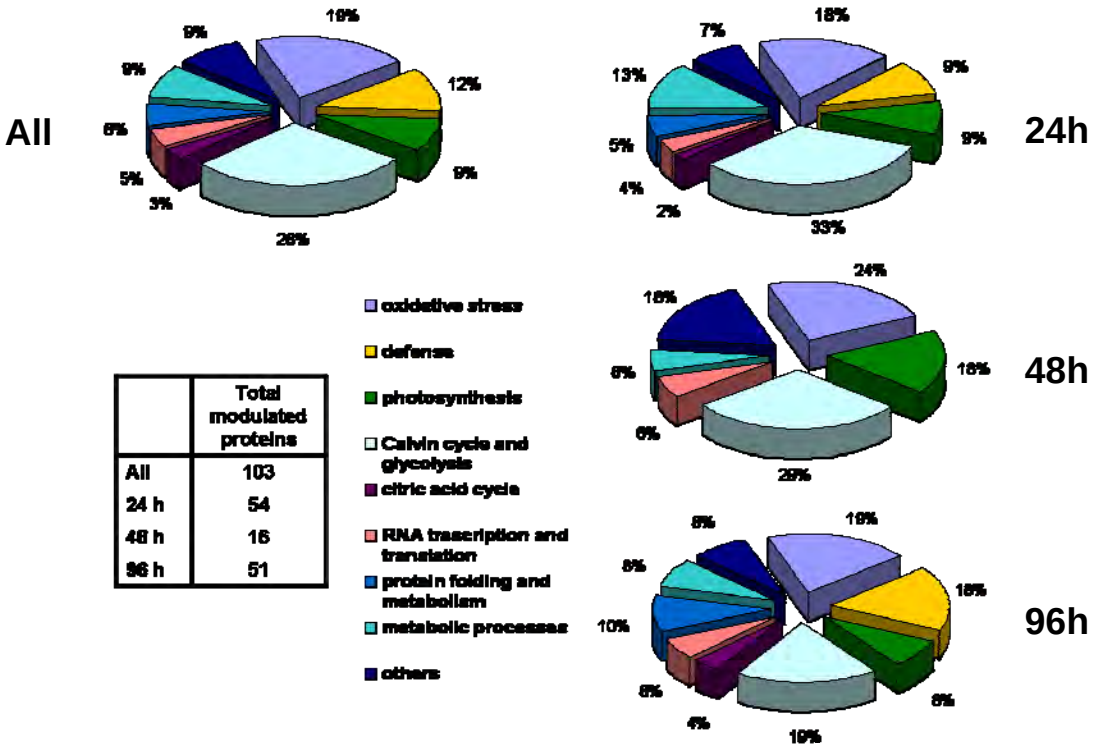

Fig. 5. Proteins GO Categorization. Distribution of the identified proteins according to the biological process in which they are involved. The pie-chart on the left represents the distribution of all the 103 identified proteins, whereas on the right there are the distributions of the identified modulated proteins after the different treatments (24, 48, and 96 hpi).

4.2.4 DISCUSSION

In this work, the *V. vinifera-P. viticola* interaction has been investigated by proteomic analysis on susceptible *V. vinifera* cultivar (Pinot Noir) leaves, collected at 24, 48, and 96 h post-infection (hpi). It is worth noting that, at present, no proteomic studies have been published on the interaction between V*itis* spp. and *P. viticola.* It is widely demonstrated that the majority of plant defence mechanisms, such as basal resistance and the hypersensitive response, are early induced during infection via a complex network of signals that occur after perception of the pathogen by host cells [16, 17]. This proteomic study aimed to gain information about proteins modulated during the first stages of *P. viticola* infection (24 and 48 hpi) and also during the later stages (96 hpi) when the pathogen spreads through the tissues and sporulates. The observation of a reduced number of modulated proteins at 48 hpi, in comparison to the other two collection times, is difficult to explain, also in consideration of the almost complete lack of information about the proteomic changes in plant oomycete interactions soon after

infection. However, some kinds of defence responses follow this trend in susceptible plants, with an early, unspecific activation of defence response, a subsequent decrease, possibly due to pathogen suppression of defences, and a late re-activation, which is not effective enough to stop pathogen spread [18]. This explanation is supported by the fact that no defence-related proteins are regulated after 48 hpi. On the other side, the enhanced protein modulation after 96 hpi may be the result of the following successful invasion of oomycete.

Among the differentially expressed proteins during the *V. vinifera-P. viticola* challenge, several polypeptides related to the photosynthesis category, more often down-regulated in comparison to healthy controls, were found. Such examples include ATP synthase CF1 α chain and ε chain, photosystem I reaction center subunit N, PsbP. It is well established that plants infected with biotrophic fungal pathogens, such as powdery mildews and rusts, reduce their photosynthetic rates, probably as a result of increased invertase activity which causes carbohydrate accumulation [19]. This decrease is measured at late time points after infection, as a reduced chlorophill content and CO_2 assimilation by fluorimetric techniques [20] and is associated to chlorosis of infected tissues. However, similar events have never been investigated in early time points as the ones that were considered in this work. Transcriptional down-regulation of genes related to photosynthesis was reported in downy mildew-infected grapevine at the oil spot stage [21]. These data would support the hypothesys that down-regulation of photosynthesis-related genes and proteins is a very early event after infection in susceptible species, occurring long before the onset of visible chlorosis symptoms. Similarly, among down-regulated proteins were ATP synthase CF1 α chain (SSP 19) and ε chain (22 and 63). ATP synthase catalyzes the formation of ATP from ADP, and Pi in the presence of proton gradient across the thylakoid membrane forming during electron transport [22]. The down-regulation after 48 hpi and 96 hpi of ATP synthase, indicates a progressive inhibition of photosynthesis-mediated energy production. On the other hand, two up-regulated proteins related to the photosystem II (PSII) were identified, namely, HCF136 protein (SSP 8) and extrinsic protein PsbP (SSP 26) respectively. Both these proteins are involved in PSII assembly/stability [23, 24]. Their up-regulation may represent a response to the photosystem damage.

Also the proteins associated with primary carbohydrate metabolism, and regulation, presented altered expression patterns following *P. viticola* infection. It is possible to speculate that the up-regulation of these proteins is a consequence of invasion while the pathogen attempts to acquire resources from the plant, for growth and survival, in a compatible interaction. Accordingly, a general increase of these proteins has been reported in infected *V. vinifera*. Many of the identified proteins belong to the glycolysis and gluconeogenesis/Calvin cycle category. Most of them are related to the major CO_2^- fixing enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO), which plays a bifunctional role, as a carboxylase (for mediating photosynthetic CO_2 assimilation) and as an oxygenase (for catalyzing the first step of the photorespiratory pathway). Increased respiration was also frequently reported in different pathosystems, as a means to obtain new energy to counteract pathogen infection [4]. In this study 7 proteins have been identified (SSP 17, 27, 59, 72, 73, 77, and 98) corresponding to the RuBisCO large subunit (LSU) and 2 proteins (SSP 21and 93) corresponding to the RuBisCo small subunit (SSU). Interestingly, 6 out of 7 LSU isoforms showed increased expression levels, while the two SSU isoforms showed a decrease in the leaves after infection. This particular correlation has never been found before; nevertheless in a recent work on *V. vinifera* cv. Pinot Noir it has been demonstrated that decreased photosynthesis was coupled with repression of RuBisCo SSU genes following fungicide stress [25]. Moreover it has been shown that (SSU) mRNA and protein decreased considerably in potato *(Solanum tuberosum*) leaves upon infection with the pathogenic fungus, *Phytophthora infestans,* or upon treatment with an elicitor preparation from the fungal culture fluid [26]. These data are in agreement with the findings here reported showing both a general decrease of photosynthesis-related proteins and a down-regulation of RuBisCo SSU. Six identified proteins correspond to glyceraldehyde-3-phosphate dehydrogenase (GADPH), (SSP 15, 45, 53, 54, 78, and 79), and all these protein isoforms were found to be up-regulated after 24 hpi or 96 hpi. Cytosolic GAPDH is involved in the second phase of glycolysis, catalysing the conversion of D-glyceraldehyde 3-phosphate into 3-phospho-D-glyceroyl phosphate. It has been suggested that these proteins may represent ‘stress-induced metabolic response’ proteins [27]. The possibility that the increase in GAPDH abundance is due simply to an increase in metabolic turnover of the protein, and not a specific stress response, should

not be overlooked. In addition, GAPDH may have a role in plant signalling pathways, accordingly with increasing evidence indicating that plant enzymes may be multi-functional [28]. For example, GAPDH has been demonstrated to be an inhibitory target of hydrogen peroxide in *A. thaliana* with potential role in mediating reactive oxygen species (ROS) signalling [29]. Interestingly, mRNA and enzyme levels of GAPDH are also induced by abscisic acid treatment [30].

Another class of identified proteins was related to oxidative stress. The plasma membrane of plant cells produces ROS, H_2O_2 and O_2^-, in response to both biotic and abiotic stimuli that play an important role in plant–pathogen interactions [31]. High intracellular levels of H_2O_2 cause the activation of plant cell death and defense mechanisms during pathogen invasion [32]. Nevertheless the growth and reproduction of plant cells require a balance between the generation of reactive molecules and the capacity of antioxidant systems to eliminate them [33]. Production of ROS in a cell, because of their reactivity, poses significant harm to various cellular components [34]. Detoxification of ROS is mediated by cellular antioxidant systems, and many proteins with known antioxidant functions were identified in the *V. vinifera-P. viticola* challenge here studied. It should be noted that an impaired modulation of pro- and anti-oxidant enzymes expression may lead to ROS levels too low for an efficient response to pathogen invasion [34, 35]. Interestingly, all the identified proteins involved in oxidative stress were found to be up-regulated in infected leaves. Among them, catalase (CAT), (SSP 57, 86, 87, and 89) was highly induced at 24 hpi with *P. viticola.* CAT, which converts H_2O_2 to water and O_2, is a key enzyme of H_2O_2-detoxification system in chloroplasts and cytosol. The significant increase in levels of CAT during the compatible interaction beetwen the susceptible *V. vinifera* and the pathogen might lead to a decrease in H_2O_2 levels. This may result in reduced plant cell death and diminished defense mechanisms in this species upon pathogen infection. Another up-regulated protein with antioxidant functions is a secretory peroxidase (SSP 97). H_2O_2 plays a crucial role in peroxidase-catalyzed cross-linking reactions during lignification, which reinforces cell walls [36, 37]. It has been reported that peroxidase expression correlates with resistance in many plant species (including barley, tobacco, wheat, and rice). In general, it has been proposed that peroxidases are involved in controlling the invasion of pathogen by the production of reactive oxygen species and by the reinforcement of

physical barriers that prevent pathogen penetration of cell walls. However, it has been also reported that rice peroxidase protein, for example, was not differentially expressed between the incompatible and compatible interactions during rice-rice blast fungus infection, suggesting that some peroxidases may be involved in a general defense response [38]. Thus, it is possible to speculate that the up-regulation of this peroxidase observed in susceptible Pinot Noir challenged by *P. viticola* may represent a not sufficient effort to prevent the infection.

Several proteins regulated after *P. viticola* infection were identified as members of the pathogenesis related proteins (PR proteins). PR proteins are produced both in susceptible and in resistant plants following infection, but their accumulation is generally much faster and stronger in resistance. They are presently classified in 17 biochemical families [11] and often display a direct antimicrobial activity, although the exact mode of action is not always known. Interestingly, in the compatible interaction investigated, expression of the PR proteins was found to be mainly modulated at 96 h after inoculation, too late to contribute to effectively stop the pathogen, which has already fully invaded the infected tissues. Among them diverse members of PR-10 family (SSP 23, 25, 34, 35, 48, 49, 50, and 61) were identified, most of which were found to be up-regulated in infected leaves. The PR-10s are a ubiquitous class of intracellular defense-related proteins (in contrast to most PR proteins which have an extracellular nature). A “ribonuclease-like function” has been attributed to PR-10s, due to sequence homology with ginseng ribonucleases [39, 40]. PR-10s also share amino acid sequence similarity to pollen allergens of trees [41] and the major food allergen of celery [42]. Most plant PR-10 proteins, such as sorghum PR-10 [43], barley PR-10 [44], and asparagus PR-10 [45], are activated upon pathogen attack or after treatment with fungal elicitors. In other plants, PR-10s are induced by drought [46], salt stress [47], and the plant hormones abscisic acid and methyl jasmonate [47]. It would seem that some PR-10 proteins may play a specific role in response to biotic stress such as pathogen invasion, perhaps providing opportunities for the development of pathogen-tolerant crop species. The delayed activation of PR-10s together with a possibly not sufficient expression level of these proteins observed in Pinot Noir upon infection by *P. viticola*, fits with the susceptibility of this species to the pathogen. Basic endochitinase (SSP 99) was another identified PR-protein, regulated after pathogen infection. Chitinases are

hydrolytic enzymes that inhibit the growth of many fungi *in vitro* by hydrolyzing the chitin of fungal cell walls. Members of these families are induced when exposed to pathogens and display antifungal activity [48]. However, this enzyme may not represent an efficient resistance response against *P. viticola*, which is an oomycete organism with a cell wall mainly constituited by glucanes rather than chitin.

Among the other differentially expressed proteins there were chaperonin 21 (SSP 28), a chaperonin GroEL-related protein (SSP 20 and 98), peptidyl-prolyl cis-trans isomerase 1 (SSP 74 and 92), and proteasome alpha type 7 (SSP 76). All these proteins showed increased levels in Pinot leaves infected with the pathogen. The role of these proteins in protein folding is amply demonstrated; however, their role in response to pathogen invasion has never been investigated. In particular, it would be interesting to deepen the study about this folding-related proteins to evaluate their possible role in pathogen susceptibility or in defence.

4.2.5 CONCLUSIONS

In this study, differentially expressed protein profiles of susceptible *V. vinifera* at different time points after *P. viticola* infection have been characterized. Comparative proteomics is a powerful tool to identify changes in the protein expression of the pathosystem. In particular, 72 nonredundant regulated proteins were identified using 2D gel proteomic approach coupled with nano-RP-HPLC-ESI-LC-MS/MS. A general decrese in photosynthesis-related proteins expression and a concomitant overexpression of carbohydrate metabolism enzymes, oxidative stress-related proteins, and PR proteins have been reported. To our knowledge, the data presented here represent the first application of proteomics to investigate a grapevine-pathogen interaction. The data thus obtained indicate that the 2-D PAGE based protocol is a sensitive and accurate method to identify protein pattern changes and can provide fascinating contributions to the understanding of complex mechanisms, such as plant-pathogen interactions. It would be of great interest to compare these results with those concerning defense response on a resistant cultivar/species at different time points after *P. viticola* infection. However, the proteins identified in this work provide a first insight into the grapevine responses to oomycete aggression.

4.3 REFERENCES

[1] N. Sanabria, D. Goring, T. Nürnberger, I. Dubery, New Phytol. 178 (2008) 503-514.

[2] J.D. Jones, J.L Dangl, Nature 444 (2006) 323-329.

[3] B.B. Buchanan, W. Gruissem, R.L. Jones, American Society of Plant Physiologists, Rockville, 2000.

[4] G.N. Agrios, Plant Pathology, fifth edition, Academic Press, San Diego, 2005.

[5] M. Moroldo, S. Paillard, R. Marconi, L. Fabrice, A. Canaguier, C. Cruaud, V. De Berardinis, C. Guichard, V. Brunaud, I. Le Clainche, S. Scalabrin, R. Testolin, G. Di Gaspero, M. Morgante, A.F. Adam-Blondon, BMC Plant Biol. 8 (2008).

[6] L.J. Grenville-Briggs, P. van West, Adv. Appl. Microbiol. 57 (2005) 217-243.

[7] J.H. Chang, A.K. Goel, S.R. Grant, J.L. Dangl, Curr. Op. Microbiol. 7 (2004) 11-18.

[8] S. Kamoun, Annu. Rev. Phytopathol. 44 (2006) 41-60.

[9] A. Kortekamp, E. Zyprian, J. Plant Physiol. 160 (2003) 1393-1400.

[10] P. Jeandet, A.C. Douillet-Breuil, R. Bessis, S. Debord, M. Sbaghi, M. Adrian, J. Agric. Food Chem. 50 (2002) 2731-2741.

[11] L.C. van Loon, M. Rep, C.M.J. Pieterse, Annu. Rev. Phytopathol. 44 (2006) 135-162.

[12] N.D. Padliya, B. Cooper, Proteomics 6 (2006) 4069-4075.

[13] T. Ebstrup, G. Saalbach, H. Egsgaard, Proteomics 5 (2005) 2839-2848.

[14] L.J. Grenville-Briggs, A.O. Avrova, C.R. Bruce, A. Williams, S.C. Whisson, P.R.J. Birch, P. van West, Fungal Genet. Biol. 42 (2005) 244-256.

[15] H.J. Mitchell, K.A. Kovac, A.R. Hardham, Mycological Research 106 (2002) 1211-1223.

[16] J.L. Dangl, J.D.G. Jones, Nature 411 (2001) 826-833.

[17] P.J.G.M. de Wit, Cell. Mol. Life Sci. 64 (2007) 2726-2732.

[18] K.H. Sohn, R. Lei, A. Nemri, J.D. Jones, Plant Cell. Dec;19 (2007) 4077-4090.

[19] J.D. Scholes, P.J. Lee, P. Horton, D.H. Lewis, New Phytol. 126 (1994) 213-222.

[20] S.A. Rolfe, J.D. Scholes, New Phytologist 83 (1995) 499-508.

[21] M. Polesani, F. Desario, A. Ferrarini, A. Zamboni, M. Pezzotti, A. Kortekamp, Annalisa Polverari, BMC Genomics 9 (2008).

[22] W.G. Hopkins, N.P.A. Hüner, Introduction to plant physiology, New York, Wiley, 2004.

[23] H. Plucken, B. Muller, D. Grohmann, P. Westhoff, L.A. Eichacker, FEBS Lett. (2002) 532 85-90.

[24] X. Yi, S.R. Hargett, H. Liu, L.K. Frankel, T.M. Bricker, J. Biol. Chem. 282 (2007) 24833-24841.

[25] A.N. Petit, G. Wojnarowiez, M.L. Panon, F. Baillieul, C. Clément, F. Fontaine, N. Vaillant-Gaveau, Planta 229 (2009) 497-506.

[26] E. Kombrink, K. Hahlbrock, Planta 181 (1990) 216-219.

[27] A.M. Laxalt, R.O. Cassia, P.M. Sanllorenti, E.A. Madrid, A.B. Andreu, G.R. Daleo, R.D. Conde, L. Lamattina, Plant Mol. Biol. 30 (1996) 961-972.

[28] B. Moore, Trends Plant Sci. 9 (2004) 221-228.

[29] J.T. Hancock, D. Henson, M. Nyirenda, R. Desikan, J. Harrison, M. Lewis, J. Hughes, S.J. Neill, Plant Physiol. Biochem. 43 (2005) 828-835.

[30] R. Velasco, F. Salamini, D. Bartels, Plant Mol. Biol. 26 (1994) 541-546.

[31] A. Levine, R. Tenhaken, R. Dixon, C. Lamb, Cell 79 (1994) 583-593.

[32] H. Takahashi, Z. Chen, H. Du, Y. Liu, D.F. Klessig, Plant J. 11 (1997) 993-1005.

[33] R. Mittler, Trends Plant Sci. 7 (2002) 405-10.

[34] R. Mittler, S. Vanderauwera, M. Gollery, F. Van Breusegem, Trends in Plant Sci. 9 (2004) 490-498.

[35] J.F. Dat, R. Pellinen, B. Van De Cotte, C. Langebartels, J. Kangasjärvi, D. Inzé, F. Van Breusegem, Plant J. 33 (2003) 621-632.

[36] P.D. Olson, J.E. Varner, Plant J. 4 (1993) 887-892.

[37] H. Thordal-Christensen, Z. Zhang, Y. Wei, D.B. Collinge, Plant J. 11 (1997) 1187-1194.

[38] S.T. Kim, S.G. Kim, D.H. Hwang, S.Y. Kang, H.J. Kim, B.H. Lee, J.J. Lee, K.Y. Kang, Proteomics 4 (2004) 3569-3578.

[39] L.C. van Loon, W.S. Pierpoint, T. Boller, V. Conejero, Plant Mol. Biol. Rep. 12 (1994) 254-264.

[40] G.P. Moiseyeva, L.I. Fedoreyevab, Y.N. Zhuravlevc, E. Yasnetskayac, P.A. Jekeld, J.J. Beintemad, FEBS Lett., 407 (1997) 207-210.

[41] H. Breiteneder, K. Pettenburger, A. Bito, R. Valenta, D. Kraft, H. Rumpold, O. Scheiner, M. Breitenbach, EMBO J. 8 (1989) 1935-1938.

[42] H. Breiteneder, K. Hoffmann-Sommergruber, G. O'Riordain, M. Susani, H. Ahorn, C. Ebner, D. Kraft, O. Scheiner, Eur. J. Biochem. 233 (1995) 484-489.

[43] S.C.C. Lo, J.D. Hipskind, R.L. Nicholson, Mol. Plant-Microbe Interact. 12 (1999) 479-489.

[44] S. Steiner-Lange, A. Fischer, A. Boettcher, I. Rouhara, H. Liedgens, E. Schmelzer, Knogge W., Mol. Plant-Microbe Interact. 16 (2003) 893-902.

[45] S.A.J. Warner, R. Scott, J. Draper, Plant Mol. Biol. 19 (1992) 555-561.

[46] L. Pnueli, E. Hallak-Herr, M. Rozenberg, M. Cohen, P. Goloubinoff, A. Kaplan, R. Mittler, Plant J. 31 (2002) 319-330.

[47] A. Moons, E. Prinsen, G. Bauw, M. Van Montagu, Plant Cell 9 (1997) 2243-2259.

[48] A. Kasprzewska, Cell. Mol. Biol. Lett. 8 (2003) 809-824.

PUBLICATIONS

DOI 10.1002/pmic.200500294 *Proteomics* 2006, *6,* S256–S261

RESEARCH ARTICLE

Searching for markers of Creutzfeldt-Jakob disease in cerebrospinal fluid by two-dimensional mapping

Chiara Piubelli[1], Michele Fiorini[1]*, Gianluigi Zanusso[1], Alberto Milli[2], Elisa Fasoli[1], Salvatore Monaco[1] and Pier Giorgio Righetti[3]*

[1] Department of Neurological and Visual Sciences, Section of Neurology, University of Verona, Verona, Italy
[2] Department of Agricultural and Industrial Biotechnologies, University of Verona, Verona, Italy
[3] Department of Chemistry, Materials and Engineering Chemistry, Milan's Polytechnic, Milan, Italy

Differential proteomic analysis has been performed on the cerebrospinal fluid (CSF) of six healthy and six patients suffering form sporadic Creutzfeldt-Jakob disease (sCJD), age- and sex-matched, after immuno-subtraction of albumin and immunoglobulins. These maps have revealed 28 polypeptide chains differentially modulated in the sCJD samples, of which 10 appeared to be up-regulated, the remaining 18 being down-regulated. Among those, 13 could be identified upon digestion and MALDI-TOF, MS analysis. In addition, the strong modulation of cystatin C was also confirmed by immunoblot analysis and the highly altered level of the 14-3-3 proteins that escaped detection by 2-D mapping, could be assessed by Western blots and immuno-detection of monomeric and homo- and hetero-dimeric 14-3-3 isotypes. In search for a panel of potential markers for sCJD, we highlight cystatin C, 14-3-3 proteins, transferrin, ubiquitin, Apoliprotein J and perhaps some of the still unidentified, but strongly modulated polypeptide chains detected in the differential map.

Received: May 2, 2005
Revised: August 13, 2005
Accepted: August 22, 2005

Keywords:
Cerebrospinal fluid / Creutzfeldt-Jakob disease / Cystatin C / 14-3-3 Proteins / Proteome analysis

1 Introduction

Sporadic Creutzfeldt-Jakob disease (sCJD), the most common human prion disorder, displays a wide clinical and pathological variability. This variability is likely determined by a site of methionine/valine (M/V) polymorphism, at the codon 129 in the prion protein gene, and by the presence of two major conformers (type 1 and type 2) of the pathological prion protein (PrP^{Sc}) [1].

Correspondence: Professor Pier Giorgio Righetti, Politecnico di Milano, Department of Chemistry, Materials and Engineering Chemistry "Giulio Natta", Via Mancinelli 7, Milano 20131, Italy
E-mail: piergiorgio.righetti@polimi.it
Fax: +39-02-23993080

Abbreviations: ACT, α_1-antichymotrypsin; **AD**, Alzheimer disease; **CSF**, cerebrospinal fluid; **PrP^{Sc}**, pathological prion protein; **sCJD**, sporadic Creutzfeldt-Jakob disease

Although a definite diagnosis requires a post mortem neuropathological and biochemical examination of brain tissue, the *in vivo* diagnosis of possible or probable sCJD can be obtained mainly by clinical observations and supportive laboratory tests. Recently, the presence of 14-3-3 proteins in the cerebrospinal fluid (CSF) has been shown to increase the diagnostic sensitivity in patients with suspected prion diseases. Because of its availability and its direct association with the brain, CSF is commonly analyzed for the *in vivo* diagnosis.

However, due to the detection of 14-3-3 proteins in patients with neurological disorders other than sCJD, to increase the sensitivity and specificity of 14-3-3 assay, we have recently determined the 14-3-3 isotype composition in association with CSF τ protein levels [2]. In addition, we have been encouraged to search for more specific markers to be used in the clinical diagnosis of sCJD and other neurodegenerative diseases.

* These authors contributed equally to this work.

In the present work, we studied CSF samples from sCJD and age-matched control subjects by a 2-D proteomic approach for the identification of proteins differentially expressed in sCJD.

2 Materials and methods

2.1 Patients and sample preparation

We studied CSF samples from subjects with definite sCJD and from neurologically and cognitively normal individuals, undergoing lumbar puncture for diagnostic or surgical purposes, according to the guidelines of the local ethical committee. For each CSF sample, a written informed consent was obtained from the patients or patients' next of kin.

Control group consisted of six non-neurological male patients, age from 53 to 76 years. sCJD group consisted of six male patients, PrP^{Sc} type 1; codon 129 genotype: M/M; age: from 56 to 76 years. A 2-mL sample of CSF containing 1 mM Pefablock protease inhibitor (Sigma-Aldrich) from each subject was treated with the Albumin and IgG Removal Kit (Amersham) (300 μL of resin per 1 mL of CSF). CSF samples were precipitated with acetone:methanol (8:1 v:v) at −20°C overnight and centrifuged at 27 000 × *g* for 30 min. Pellets were resuspended in 2-D solubilizing solution (7 M urea, 2 M thiourea, 2% CHAPS, 40 mM Tris, 1% Pharmalytes 3.5–10, 3 mM tributylphosphine) with the addition of a protease inhibitor cocktail (Complete, Roche Diagnostic). An aliquot of each sample was dialyzed and used for protein concentration determination by using the Bio-Rad Dc protein assay, according to the manufacturer's instructions.

2.2 2-D proteomic analysis and MALDI-TOF MS

CSF samples, each containing 0.6 mg of total protein, were processed for 2-DE in non-linear pH 3–10 IPG strips and in 8–18% acrylamide gradient SDS-PAGE [3]. The 2-D gels were stained by SYPRO Ruby fluorescent staining. Differential analysis was performed on 12 gels generated from 12 samples. The 2-D gels were land marked and matched in a matchset by the PDQuest software (Bio-Rad). Spot intensities were normalized in each gel and a Student's *t*-test ($p < 0.05$) was adopted for evaluating statistically significant differences between the control and the CJD group [4]. The differentially expressed spots were excised by means of a spot cutter robot (Bio-Rad); protein digestion and MS analysis was performed as in [5].

2.3 Immunoblot analysis

2.3.1 Cystatin C

SDS-PAGE was performed on 2.5 μL of CSF of patients with sCJD and control subjects. The Western blot was incubated with anti-cystatin C polyclonal antibody (Upstate) at a 1:1000 dilution. Detection was implemented by using an ECL system (Amersham).

2.3.2 14-3-3 proteins

Detection of 14-3-3 proteins was performed after SDS-PAGE of 25 μL of CSF from sCJD and control subjects. The Western blot was incubated with anti-14-3-3 K-19 pan polyclonal antibody (SC-629, Santa Cruz Biotechnology) at a 1:500 dilution and revealed by using an ECL system (Amersham). All samples positive to the pan 14-3-3 antibody were also tested with antibodies recognizing specific isoforms (all from Santa Cruz Biotechnology), including anti-β (SC-17288), anti-γ (SC-731), anti-ε (SC-1020), anti-σ (SC-7683), anti-η (SC-17286), anti-τ (SC-732) and anti-ζ (SC-1019) at 1:500 dilution. By the latter technique, it was possible to detect strong increments in the γ and ε isotypes via 2-D mapping.

3 Results and discussion

3.1 General remarks

Figure 1A shows the reference map resulting from the matchset of 12 2-D maps from 12 CSF samples analyzed (left panel). The central and right panels show typical 2-D maps of a sCJD and a control sample, respectively. The significant spots resulting from statistical analysis are indicated in all three maps. Twenty-eight spots were found to be differentially expressed in the CJD group compared to the control group; of these, 10 appeared to be up-regulated, the remaining 18 being down-regulated. The spots were eluted, digested and subjected to MALDI-TOF MS for PMF identification analysis. About half of them (13) could be identified, as shown in Table 1. Figure 1B shows up-regulated CSF proteins, as identified by MALDI analysis, in addition to their relative amounts as compared to controls. To verify the results of differential proteomic analysis, we attempted detection by immunoblots of some of the altered polypeptide chains. Figure 2 gives an example of such a Western blot identification, in the case of cystatin C, a protein found to be extremely overexpressed (>30-fold) in the CSF of sCJD patients. It can be appreciated that the same trend is exhibited also in the immunoblot, with a markedly more intense spots in the three patients (representative of all studied subjects), as compared to the control. Additionally, although not quite detectable in the 2-D maps, we have attempted detection of the 14-3-3 proteins by SDS-PAGE immunoblot revelation. As shown in Fig. 3A, although 14-3-3 is barely discernible in controls, we obtain a very strong signal in all sCJD patients, of which three samples are shown in the blot. By combining 2-D separations of 100 μL CSF samples and Western blots with different anti-14-3-3 antibodies, we were able to detect traces of monomeric and dimeric 14-3-3 proteins in normal controls. On the other hand, the study of CSF

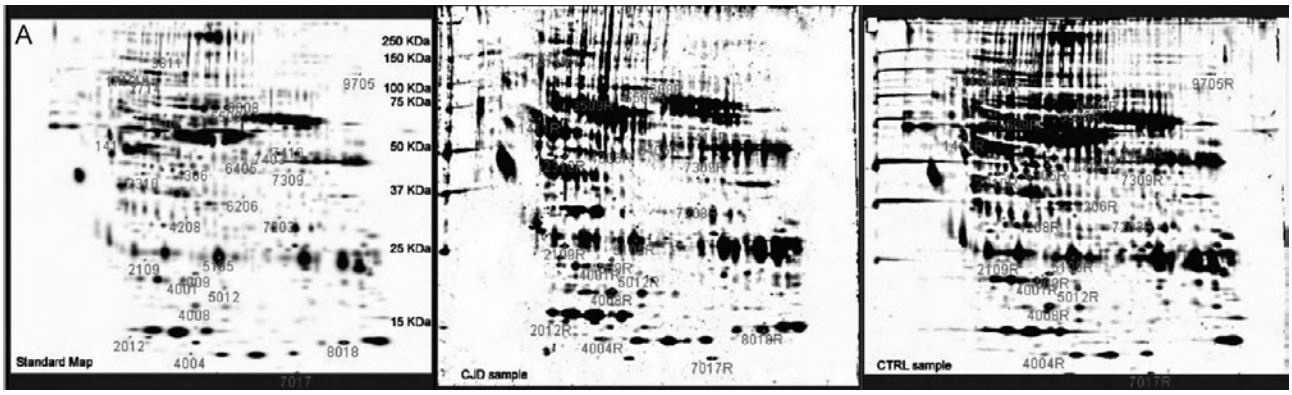

Figure 1. (A) Representative 2-D maps of CSF from: (a) control subjects (right panel), (b) sCJD patients (central panel) and (c) match-set synthetic map. In all cases, the numbers indicate the spots up- or down-regulated. (B) Relative normalized amounts of identified proteins showing increased expression in the CSF of sCJD subjects (black bars), in comparison to age-matched normal controls (grey bars). Averages and standard deviations were obtained from six sCJD cases and six controls.

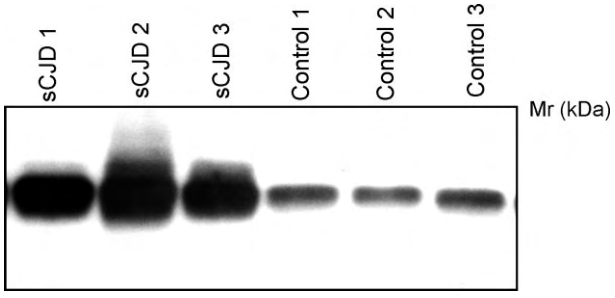

Figure 2. Detection of cystatin C in 25 μL of CSF of patients with sCJD and control subjects after immunoblot from 1-D SDS-PAGE. Blot revealed with an anti-cystatin C polyclonal antibody at a 1:1000 dilution followed by an ECL system.

samples from sCJD subjects disclosed very high signals for the gamma and epsilon isotypes, in addition to dimeric species (Fig. 3B). The dimeric nature of these higher molecular mass species, as well as their presence in all control and pathological cases, was confirmed with the use of isotype-specific antibodies (for each CSF sample at least eight 2-D Western blots were obtained), antibody-omission studies, and, more importantly, by comparison with SDS-PAGE and 2-D immunoblots of human recombinant 14-3-3 isoforms and a mixture thereof (data not shown). Previous studies have shown that the formation of stable 14-3-3 molecules is due to interfacing of multiple hydrophobic regions between monomeric forms. In this context, the occurrence of strong hydrophobic interactions is also suggested by resistance of these high molecular mass species to reducing and chaotropic agents. The significance of some of these findings, in relation to the altered expression of some of the proteins listed in Table 1, will be discussed below.

3.2 Transferrin (+5-fold)

We have been unable to find data on modulation of this protein in CSF, but at least one report [6] described a fivefold upregulation (the same level found by us) in sera of patients with Alzheimer disease (AD). In general, the trend of regulation in sera is similar to the one found in CSF. The role of transferrin, in brain iron homeostasis, has been recently reviewed by Moos [7].

Table 1. Identification of CSF proteins differentially expressed in CSF from sporadic Creutzfeldt-Jakob patients

Spot number	Trend sCJD *vs.* controls	Fold changes	Statistical significance ($p < 0.05$)	Protein identification	MASCOT score	Matched peptides	Coverage	Swiss-Prot ID
1411	Up-regulation	3	0.021	α1-Antichymotrypsin	151	18	55	P01011
1713	Down-regulation	2.7	0.038					
2012	Up-regulation	Very high	0.003					
2109	Up-regulation	6	0.015	Apolipoprotein A1	170	21	72	P02647
2310	Up-regulation	3.6	0.011	Apolipoprotein A4	130	25	66	P06727
2714	Down-regulation	Very high	0.002					
2716	Up-regulation	14.3	0.006					
3509	Down-regulation	6.4	0.041					
3811	Down-regulation	5	0.025					
4001	Down-regulation	1.5	0.012					
4004	Up-regulation	4	0.050					
4008	Up-regulation	3.3	0.021	Hp2-α haptoglobin	117	18	45	P00738
4009	Down-regulation	2.4	0.011					
4208	Down-regulation	2.4	0.037	Apolipoprotein J	58	14	35	P10909
4306	Down-regulation	2.5	0.030	Fibrinogen γ-chain	87	13	40	P02679
5012	Down-regulation	2	0.014					
5105	Down-regulation	2	0.003	Prostaglandin D2 synthase/IgG light chain	88	8	46	P41222
5509	Down-regulation	2	0.019	Gelsolin	98	33	40	P06396
6206	Down-regulation	Very high	0.001	Fibrin β	77	19	48	
6405	Down-regulation	8	0.040					
6608	Down-regulation	1.6	0.011	Complement factor B/3a	74	23	36	P00751
7017	Up-regulation	13	0.047	Ubiquitin	65	5	56	P62988
7203	Down-regulation	4.3	0.012					
7309	Up-regulation	2.6	0.012					
7403	Down-regulation	2.6	0.033					
7416	Up-regulation	5	0.020	Transferrin	135	36	50	P02787
8018	Up-regulation	Very high	0.001	Cystatin C	74	7	51	P01034
9705	Down-regulation	9.4	0.007					

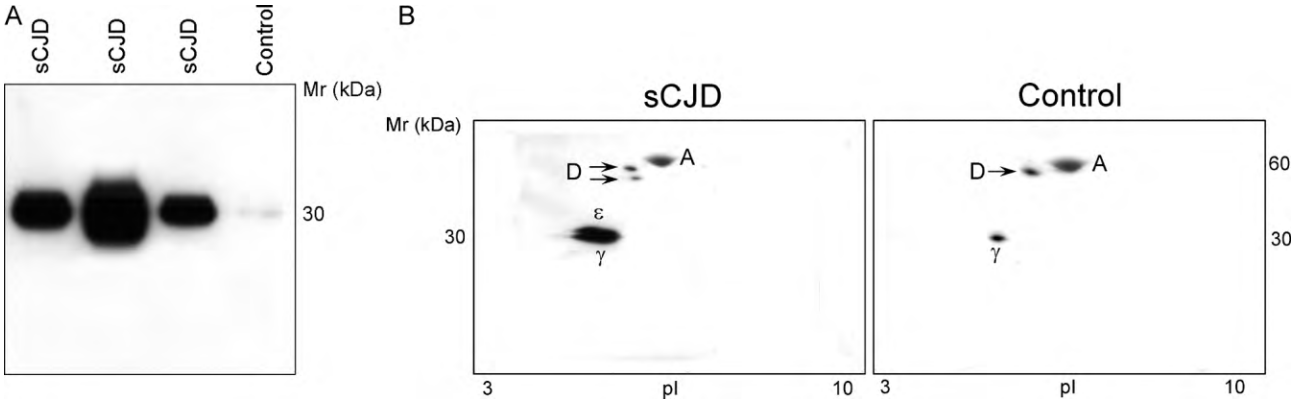

Figure 3. (A) Detection of the 14-3-3 protein in 25 μL of CSF of patients with sCJD and control subjects after immunoblot from 1-D SDS-PAGE. Blot revealed with anti-14-3-3 K-19 pan polyclonal antibody, at a 1:500 dilution, followed by an ECL system. (B) Detection of 14-3-3 in 100 μL of CSF from a patient with sCJD and a control subject after immunoblot from 2-D-PAGE. Blot revealed with anti-14-3-3 K-19 pan polyclonal antibody at a 1:500 dilution, followed by an ECL system. The presence of dimeric 14-3-3 species (D) is seen in both preparations. A denotes albumin.

3.3 α_1-Antichymotrypsin (ACT; +3-fold)

Licastro *et al.* [8] have reported an increase of ACT in the plasma of AD patients, although they question if these increments could be specific for AD or might be a general manifestation of a progressive brain inflammation, as engendered by AD. Elevated levels of α_1-Antichymotrypsin (ACT) in plasma of AD subjects have also been described by Wang *et al.* [9]. An interesting case is the one reported by DeKosky *et al.* [10] where the elevated levels of ACT in both sera and CSF of AD patients were found. Interestingly, Wang and colleagues observe that such increments of ACT are to be found in AD, but not in dementias other than AD, and that ACT keeps increasing with progression of AD dementia. Their conclusions, though, are that ACT is not useful as a diagnostic biomarker for AD.

3.4 Apolipoprotein J (Apo-J) (clusterin; −2.5-fold)

The 2-D mapping of CSF from AD patients has shown a major modulation of eight proteins, including Apo-J [11]. A number of reports have focused on the altered expression of Apo-J in AD disease. According to Calero *et al.* [12], Apo-J expression is up-regulated in a wide variety of insults and may represent a defence response against local damage to neurons. However, in CJD [13] and in transmissible spongiform encephalopathies [14] affected patients, this overexpression of Apo-J in brains seems to serve a unique function, namely to participate in PrP clustering and sequestration, thus modifying PrP toxicity in CJD. Thus, if Apo-J, even if present at higher levels in affected brains, is mostly used for complexation into PrP clusters, our findings of its lowered presence in the CSF of sCDJ patients might not contradict the findings reported by Freixes *et al.* [13] and by Sasaki *et al.* [14].

3.5 Gelsolin (−2-fold)

We have been unable to find reports of variation of this protein in CSF. However, Kiuru *et al.* [15] report that, in the case of a Finnish type of familial amyloidosis, deposits of gelsolin-related amyloid are found in the brain of patients. Quite a few other reports claim that gelsolin can either form amyloidogenic fragments in brain [16] or, in any event, be strongly associated with the amyloid β-protein [17, 18], to the point of resisting dissociation in SDS in the absence of reducing agents, a finding suggesting formation of –S-S- bonds between gelsolin and the amyloid β-protein. Thus, the formation of gelsolin-amyloid β-protein complexes might lead to decreased amyloid deposition in affected brain tissues. Similarly, our present findings of gelsolin down-regulation in the CSF of sCJD subjects suggest a role for this protein in the pathogenesis of prion disorders.

3.6 Ubiquitin (+13-fold)

The overexpression of this protein in the CSF of sCJD patients is among the highest reported in this study. Ubiquitin is involved in the ATP-dependent selective degradation of cellular proteins, the maintenance of chromatin structure, the regulation of gene expression, the stress-response, and ribosome biogenesis. Several studies have reported that the malfunction of the ubiquitin-proteasome system is among the causes of neurodegeneration in Parkinson's disease, AD and other neurodegenerative disorders [19–21]. Iqbal and Grundke-Iqbal [22] have reported elevated levels of ubiquitin in both brain and cerebrospinal fluid of AD patients. Interestingly, Davidsson *et al.* [23] described a 2-D approach quite similar to the one here reported for comparing the CSF proteome of control and AD patients and found significantly altered levels of ubiquitin in the latter samples.

3.7 Cystatin C (>30-fold up-regulated)

Cystatin C is the most frequently investigated member of a super-family of protease inhibitors, since it is involved in processes such as tumor invasion and metastasis, inflammatory processes and some neurological diseases (for a review, see [24]). This is perhaps one of the most interesting markers for sCJD, since it appears to be more specific than other markers discussed so far. In fact, in an extensive study on serum and CSF in vascular and Alzheimer's dementia, both early and late onset, Kalman *et al.* [25] found that cystatin C levels were in the normal range in all groups, thus excluding the possibility that this protein could be a diagnostic marker for both types of dementia. Additionally, cystatin C has been found to be down-regulated in the CSF form patients with leptomeningeal metastasis [26] as well as from patients with inflammatory neurological diseases, such as Guillain-Barre syndrome and chronic inflammatory demyelinating polyneuropathy [27]. On the contrary, our data, as well as these of Hochstrasser's group [28], indicate this protein to be strongly up-regulated in the CSF of CJD patients. Interestingly, the latter authors find this marker essentially based on SELDI-TOF MS analysis, confirmed, as an additional check, also by a Western blot. This finding is quite interesting, considering that the SELDI technique had recently been heavily criticized by authors claiming that its sensitivity was too low (in the micromolar range) and thus could not possibly permit detection of biomarkers, typically present at the nanomolar and even lower levels [29]. Clearly, their data, as well as our own, obtained by a completely different approach, dispel this myth.

3.8 14-3-3 protein (>50-fold up-regulated)

The 14-3-3 proteins represent a family of seven isoforms (α through η) of low Mr proteins (30 kDa), which share a high degree of homology. Though ubiquitous, these proteins are particularly abundant in the nervous system, where they

represent about 1% of total cytosolic proteins [30]. The functional role of this protein is still unclear, but there are data supporting a role as activator of neurotransmitter synthesis, in particular cathecolamine and serotonin, and as a regulator of cell cycle and apoptosis [31]. Interestingly, upon immunoblot analysis, high levels of 14-3-3 were found in the CSF of sCJD patients [32], to the point at which this protein is now considered a highly sensible (96%) and specific (88%) diagnostic marker. Significantly, in our case too, although only traces of 14-3-3 could be visualized in normal subjects, its presence at very high levels could be confirmed by SDS-PAGE and 2-D in all sCJD patients here analyzed.

4 Concluding remarks

The trend, in present-day search for biomarkers, is not just to find a single polypeptide chain as an indicator of a disease, as done in the past, but rather a panel of biomarkers, whose total expression will tend towards 100% sensitivity coupled to 100% specificity, the goal in today clinical research [33]. For instance, in a recent study, Carrette *et al.* [34] identified a panel of five markers, namely cystatin C, two β_2-microglobulin isoforms, an unknown 7.7-kDa polypeptide and a 4.8-kDa VGF polypeptide that would give 66% sensitivity and 100% specificity in the diagnosis of AD in sample of CSF. In our case, given the high number of proteins modulated, and given the fact that half of them could not be identified as yet, it would be difficult to assemble the panel of markers and to decide how many of them would be necessary for unambiguous identification of sCJD. Caution is also warranted by the small cohort of investigated cases and the failure to identify a marker allowing an unambiguous differentiation between normal and pathological. Notwithstanding that, at least two proteins appear to be likely candidates: cystatin C (which might be in common, but probably at a much reduced overexpression level, considering the conflicting reports in [25] and [34] with Alzheimer's dementia) and the 14-3-3 protein. These two, supplemented with ubiquitin, transferrin, Apo-J and may be a few of those polypeptides that were strongly up-regulated but could not be identified as yet (see Table 1) could form a highly reliable panel for univocal assessment of sCJD.

Supported by the MIUR (FIRB 2001, Cod. RBNE01KJHT; PRIN, 40%, 2003 and 2005), Rome, by a grant from Cariverona Bank (year 2002), and in part by "Ricerca finalizzata 2003 "Studio dei meccanismi patogenetici delle malattie neurodegenerative per la diagnosi e lo sviluppo di approcci terapeutici" to S.M.

5 References

[1] Zanusso, G., Farinazzo, A., Prelli, F., Fiorini, M. *et al.*, *J. Biol. Chem.* 2004, *279*, 38936–38942.

[2] Zanusso, G., Fiorini, M., Farinazzo, A., Gelati, M. *et al.*, *Neurology* 2005 *64*, 1618–1625.

[3] Castagna, A., Antonioli, P., Astner, H., Hamdan, M. *et al.*, *Proteomics* 2004, *4*, 3246–3267.

[4] Marengo, E., Robotti, E., Antonucci, F., Cecconi, D. *et al.*, *Proteomics* 2005, *5*, 654–666.

[5] Piubelli, C., Cecconi, D., Astner, H., Caldara, F. *et al.*, *Proteomics* 2005, *5*,1382–1394.

[6] Yu, H. L., Chertkow, H. M., Bergman, H., Schipper, H. M., *Proteomics* 2003, *3*, 2240–2248.

[7] Moos, T., *Danish Med. Bull.* 2002, *49*, 279–301.

[8] Licastro, F., Pedrini, S., Caputo, L., Annoni, G. *et al.*, *J. Neuroimmunol.* 2000, *103*, 97–102.

[9] Wang, X., DeKosky, S. T., Ikonomovic, M. D., Kamboh, M. I., *Neurobiol. Aging.* 2002, *23*, 377–382.

[10] DeKosky, S. T., Ikonomovic, M. D., Wang, X., Farlow, M. *et al.*, *Ann. Neurol.* 2003, *53*, 81–90.

[11] Puchades, M., Hansson, S. F., Nilsson, C. L., Andreasen, N. *et al.*, *Brain Res. Mol. Brain Res.* 2003, *118*, 140–146.

[12] Calero, M., Rostagno, A., Matsubara, E., Zlokovic, B. *et al.*, *Microsc. Res. Tech.* 2000, *50*, 305–315.

[13] Freixes, M., Puig, B., Rodriguez, A., Torrejon-Escribano, B. *et al.*, *Acta Neuropathol.* 2004, *108*, 295–301.

[14] Sasaki, K., Doh-ura, K., Ironside, J. W., Iwaki, T., *Acta Neuropathol.* 2002, *103*, 199–208.

[15] Kiuru, S., Salonen, O., Haltia, M., *Ann. Neurol.* 1999, *45*, 305–311.

[16] Benyamini, H., Gunasekaran, K., Wolfson, H., Nussinov, R., *Proteins* 2003, *51*, 266–282.

[17] Ray, I., Chauhan, A., Wegiel, J., Chauhan, V. P., *Brain Res.* 2000, *853*, 344–351.

[18] Chauhan, V. P., Ray, I., Chauhan, A., Wisniewski, H. M., *Biochem. Biophys. Res. Commun.* 1999, *258*, 241–246.

[19] Mandel, S., Youdim, M.B., *Free Radical Biol. Med.* 2004, *37*, 304–317.

[20] Bossy-Wetzel, E., Schwarzenbacher, R., Lipton, S. A., *Nat. Med.* 2004, *10*, S2–9.

[21] Zhang, J., Goodlett, D. R., *Mol. Neurobiol.* 2004, *29*, 271–288.

[22] Iqbal, K., Grundke-Iqbal, I., *Intl. Psychogeriatr.* 1997, *9, Suppl. 1*, 289–296.

[23] Davidsson, P., Westman-Brinkmalm, A., Nilsson, C. L., Lidbjer, M. *et al.*, *Neuroreport* 2002, *13*, 611–615.

[24] Reed, C. H., *Br. J. Biomed. Sci.* 2000, *57*, 323–329.

[25] Kalman, J., Marki-Zay, J., Juhasz, A., Santha, A. *et al.*, *Acta Neurol. Scand.* 2000, *101*, 279–282.

[26] Nagai, A., Terashima, M., Harada, T., Shimode, K. *et al.*, *Clin. Chim. Acta* 2003, *329*, 53–60.

[27] Nagai, A., Murakawa, Y., Terashima, M., Shimode, K. *et al.*, *Neurology* 2000, *55*, 1828–1832.

[28] Sanchez, J. C., Guillaume, E., Lescuyer, P., Allard, L. *et al.*, *Proteomics* 2004, *4*, 2229–2233.

[29] Diamandis, E. P., *J. Natl. Cancer Inst.* 2004, *96*, 353–356.

[30] Boston, P. F., Jackson, P., Kynoch, P. A. M., *J. Neurochem.* 1982, *38*, 1466–1474.

[31] Berg, D., Holzmann, C., Riess, O., *Nat. Rev. Neurosci.* 2003, *4*, 752–762.

[32] Hsich, G., Kenney, K., Gibbs, C. J., Lee, K. H., Harrington, M. G., *New England J. Med.* 1996, *13*, 924–930.

[33] Liggett, W. S., Barker, P. E., Semmes, O. J., Cazares. L–H., *Disease Markers* 2004, *20*, 295–307.

[34] Carrette, O., Demalte, I., Scherl, A., Yalkinoglu, O. *et al.*, *Proteomics* 2003, *3*, 1486–1494.

European Neuropsychopharmacology (2006) **16**, 521—537

www.elsevier.com/locate/euroneuro

Proteomic analysis of rat hippocampus and frontal cortex after chronic treatment with fluoxetine or putative novel antidepressants: CRF1 and NK1 receptor antagonists

Lucia Carboni [a,*], Miriam Vighini [a,b], Chiara Piubelli [a], Laura Castelletti [a], Alberto Milli [b,c], Enrico Domenici [a]

[a] *Department of Biology, Psychiatry CEDD, GlaxoSmithKline, Verona, Italy*
[b] *Department of Agricultural and Industrial Biotechnologies, University of Verona, Italy*
[c] *Department of Computational, Analytical and Structural Science, GlaxoSmithKline, Verona, Italy*

Received 29 June 2005; received in revised form 12 January 2006; accepted 17 January 2006

KEYWORDS
2D electrophoresis;
Antidepressive agents;
Proteomics;
Neurokinin1 receptor;
Corticotropin-releasing hormone;
Serotonin uptake inhibitors

Abstract Chronic administration of antidepressants is required for their efficacy, suggesting the involvement of long-term modifications. As the impact of antidepressant treatment on the brain molecular machinery is not completely understood, we performed a proteomic analysis of rat hippocampus and frontal cortex after chronic treatment with fluoxetine, with an NK1 receptor antagonist, GR205171, and a CRF receptor 1 antagonist, DMP696. After 2D electrophoresis, protein expression levels were compared with both univariate and multivariate statistical analyses and identified by mass spectrometry. All treatments modified levels of actin isoforms, whereas both fluoxetine and GR205171 reduced synapsin II. Fluoxetine treatment increased ERK2 and NP25 and decreased vacuolar ATP synthase. After GR205171 treatment, protein disulphide isomerase A was reduced; dynamin 1 and aldose reductase increased. DMP696 modulated DRP2, pyruvate kinase, LDH and ATP synthase. Although each compound induced a specific pattern of protein modulation, data suggest that antidepressants share the ability of modulating neural plasticity.

1. Introduction

Most antidepressive agents effective in the therapy of mood disorders share the common feature of being able to increase the synaptic monoamine levels. This increase is

* Corresponding author. Psychiatry CEDD, Via A. Fleming 4, 37135 Verona, Italy. Tel.: +39 045 8218294; fax: +39 045 8218047.
E-mail address: lucia.e.carboni@gsk.com (L. Carboni).

0924-977X/$ - see front matter © 2006 Elsevier B.V. and ECNP. All rights reserved.
doi:10.1016/j.euroneuro.2006.01.007

considered only the first step in a series of unknown adaptations in brain, which are responsible for the long-term efficacy. It is widely accepted that the delay of 3–6 weeks after initiation of treatment before observation of a clinical response implies that the acute modulation of monoamine levels is not the final pathway of antidepressant action. The mechanism of action of antidepressive agents thus remains obscure (Schloss and Henn, 2004; Nemeroff and Owens, 2002; Nestler, 1998). Moreover, although available antidepressants are efficacious in relieving symptoms in a subpopulation of patients (Spigset and Martensson, 1999; Nemeroff and Owens, 2002), new agents are needed to widen the percentage of responders, with rapid onset of action and improved side effect profile. A new potential antidepressant class is represented by antagonists to the receptor 1 of the corticotropin-releasing factor (CRF), the hypothalamic peptide that controls the behavioural, endocrine autonomic and immune responses to stress (Nemeroff, 2002; Kehne and De Lombaert, 2002; Whitnall, 1993; Claes, 2004). The rationale of the approach is based on the clinical evidence of an abnormal stress response in a large percentage of depressive patients (Holsboer, 1999; Holsboer, 2003; Claes, 2004), which often show hypothalamic–pituitary–adrenal system dysfunction (Claes, 2004; Müller and Wurst, 2004). In addition, modifications in CRF levels were detected in depressive patient brains (Claes, 2004). Preclinical studies, utilising transgenic animals and tool compounds, have revealed that the effects of CRF in laboratory animals resemble depressive symptoms (Nemeroff and Owens, 2002; Nemeroff, 2002; Holsboer, 2003; Müller and Wurst, 2004; Kehne and De Lombaert, 2002), supporting this hypothesis. A CRF receptor 1 antagonist showed clinical efficacy in treating depressive patients (Zobel et al., 2000), thus providing a proof of principle for this mechanism. Although the large amount of available preclinical data, the long-term effects of pharmacological treatment with CRFR1 antagonists on molecular pathways in brain are unknown. Antagonists to the NK1 tachykinin receptor belong to another novel class of potential antidepressants (Stout et al., 2001; Rupniak, 2002; Adell, 2004). The suggestion was supported by preclinical findings, which showed an anxiogenic profile for substance P and an anxiolytic and antidepressant-like activity for NK1 antagonists in a range of behavioural models (Stout et al., 2001; Rupniak, 2002; Adell, 2004). NK1 antagonists exhibited robust antidepressant efficacy in the first clinical study, whereas further trials provided conflicting results (Adell, 2004), thus suggesting that additional research is needed to fully understand the potential of this mechanism. Both CRF1 and NK1 antagonists have been shown to be able to counteract stress-induced suppression of neurogenesis in preclinical models (Van der Hart et al., 2002; Alonso et al., 2004), suggesting that these novel classes may affect plasticity changes in the hippocampal formation in a manner similar to clinically effective antidepressants.

The study we are presenting was designed to identify changes in protein profiles induced in rat brain by chronic treatments with the antidepressant fluoxetine, with DMP696, a CRF receptor 1 antagonist, and GR205171, an NK1 antagonist. The aim was to gain insights into the long-term mechanisms of action of these classes of established or potential antidepressants. Fluoxetine is a selective inhibitor of serotonin reuptake widely used in clinical practice (Gram, 1994); DMP696 is a potent CRFR1 antagonist which shows oral bio-availability (He et al., 2000); and GR205171 is an orally active, brain-penetrant NK1 antagonist which binds with high affinity to the rat receptor (Gardner et al., 1996; Rupniak et al., 2003). The expression proteomic studies were carried out in hippocampus and frontal cortex, regions that play a key role in the pathophysiology of anxiety disorders and major depression (Posener et al., 2003; Drevets et al., 1997), and are thought to mediate the short and long-term effects of antidepressant drugs both in humans and in preclinical models (D'Sa and Duman, 2002).

2. Methods

2.1. Animals

Sprague–Dawley male rats (300 g at the beginning of the experiment, Charles River, Italy) were housed under standard laboratory conditions (22 ± 1 °C), on a 12 h light/dark cycle with lights on at 6:00 a.m. Food and water were available ad libitum. All procedures were carried out in accordance with the Italian law (art. 7, Legislative Decree no. 116, 27 January 1992), which acknowledges the European Directive 86/609/EEC, and were fully compliant with GlaxoSmithKline policy on the care and use of laboratory animals and related codes of practice.

2.2. Pharmacological treatments

All the animals arrived and were housed together; each experimental group contained eight rats. Pharmacological treatments were administered by oral gavage for 21 days following a sequential scheme in which two animals per group were added each day. Two control groups received water or 0.5% methocel, 1% dimethylsulphoxide as vehicles. Three groups treated with compounds received: fluoxetine dissolved in water, 5 mg/kg/day; GR205171 dissolved in water, 5 mg/kg twice a day; DMP696 dissolved in 0.5% methocel, 1% dimethylsulphoxide, 5 mg/kg/day. Drugs were synthesised in house. A pilot pharmacokinetic study on DMP696 in rats suggested that the dose of 5 mg/kg was likely to achieve the desired exposure, both systemically and into the brain. Pharmacokinetic studies were performed on three additional animals per treated group at the 1st, 9th and 21st days of treatment to measure the plasma concentration of the drugs. At the 14th day of treatment, the GR205171 dose was increased to two administrations of 10 mg/kg/day, since the pharmacokinetic data indicated a reduction in plasma concentrations possibly due to metabolism induction. All sacrifices were carried out between 9 and 12 a.m.

2.3. Sample preparation

The day after the last treatment the animals were guillotined and hippocampus and frontal cortex (divided into left and right halves) were dissected from brains. Tissues were frozen in dry ice and stored at − 80 °C. Each right hippocampus or frontal cortex was homogenised with a glass/teflon homogeniser at 10% (w/v) in a solubilising solution containing: 7 M urea (Sigma-Aldrich, St. Louis, MO, USA), 2 M thiourea (Fluka, Buchs, Switzerland), 40 mM Tris (Sigma-Aldrich), 3 mM tributylphosphine (Fluka), 2% (w/v) 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS; Fluka) and 1% Pharmalytes 3.5–10 (Amersham Biosciences, Uppsala, Sweden), Complete™ protease inhibitor cocktail (Roche, Basel, Switzerland). Samples were sonicated three times for 10 s on ice with an ultrasonic processor with probe (Ultrasonic 2000, Dynatech Laboratories Inc., Virginia, USA). The extract was centrifuged at

1000×*g* and the pellet was discarded. The supernatant was stored at − 80 °C until 2D electrophoresis analysis. An aliquot of this supernatant was dialysed against 1% sodium dodecyl sulphate (SDS, Sigma-Aldrich) in water and used to measure protein concentration by the BCA method (Pierce, Rockford, IL, USA) following manufacturer's instructions.

2.4. 2D electrophoresis

Since our electrophoretic apparatus can accommodate up to 24 gels in parallel, samples sharing the same vehicle were run together. Therefore, two runs were carried out with hippocampal proteins, the first one including fluoxetine, GR205171 and the respective vehicle treatment group; another one comprising DPM696 and its vehicle. Another two runs were performed following the same scheme for frontal cortex samples.

Isoelectric focusing was carried out on 17 cm immobilised pH gradient strips (3—10 non-linear pH gradient, Bio-Rad, Hercules, CA, USA). The strips were rehydrated for 16 h with 0.5 mg of proteins in 400 μl of the solubilising solution already described, with the addition of 10 mM iodoacetamide as alkylating agent (Herbert et al., 2001) and bromophenol blue. Focusing was carried out at 20 °C for 75,000 Vh at a maximum of 10,000 V in a Protean IEF Cell (Bio-Rad). Strips were then incubated with gentle shaking in an equilibration solution (6 M urea, 2% w/v SDS, 375 mM Tris pH 8.8) for 25 min. The strips were then laid on top of home-made polyacrylamide slab gels (8—18.5% T gradient, 20 × 20.5 cm and bis-acrylamide from Bio-Rad) and sealed by 0.5% w/v agarose gel prepared in running buffer and stained with bromophenol blue. An aliquot of Precision™ molecular mass marker (Bio-Rad) was loaded in parallel on each gel. Gels were run in the Protean plus Dodeca Cell apparatus (Bio-Rad) at 18 °C and 12 mA/gel overnight in 15 mM Tris, 192 mM glycine (Sigma-Aldrich), 0.1% w/v SDS. After 1 h washing in 10% v/v ethanol and 7% v/v acetic acid in water, staining was carried out overnight with Sypro Ruby fluorescent stain (Bio-Rad). Gels were destained 1 h in a 10% v/v ethanol and 7% v/v acetic acid solution and maintained in water. Images were acquired with a CCD camera on the VersaDoc imaging system (Bio-Rad).

2.5. Data analysis

Image analysis was carried out with the PDQuest software (Bio-Rad). Protein levels were evaluated as spot quantities (spot quantity = spot height$*\pi*\sigma_x*\sigma_y$, where spot height was the peak of the Gaussian representation of the spot and $\sigma_x*\sigma_y$ were the standard deviations of the Gaussian distribution of the spot in the *x* and *y* axes, respectively) for the proteins matched among gels. Spot quantities were normalised in each gel for total density in valid spots. Data were log transformed and analysed with Student's *t*-test (TT) and partial least squares analysis (PLS) with the statistics tools included in the PDQuest software. Spots giving significant results ($p<0.05$) were verified visually to exclude artefacts. When matched spots were missing in some gels, quantity values of satisfactory spots were re-checked by TT using the Prism software (GraphPad Software Inc., San Diego CA, USA).

2.6. Protein identification by peptide fingerprinting mass spectrometry

Mass spectrometry identification was performed as already described (Piubelli et al., 2005). Briefly, spots were cut, destained (2 × 10 min 200 μl washes in 50% acetonitrile (Sigma-Aldrich) (v/v), 50% 5 mM Tris pH 8.5; 1 × 10 min wash 5 mM Tris pH 8.5) and dried in a Speedvac sc110A device (Thermo Savant, NY, USA). Samples were digested with 15 μl of Sequencing Grade Modified Trypsin (Promega, Promega Corporation, Madison, WI, USA) (0.02 mg/ml) in NH_4HCO_3 buffer (40 mM, pH 8.5) at 37 °C overnight. Peptides were extracted in an ultrasonic bath for 15 min (twice 50 μl 50% acetonitrile, 50% H_2O with 1% formic acid v/v; once 50 μl of

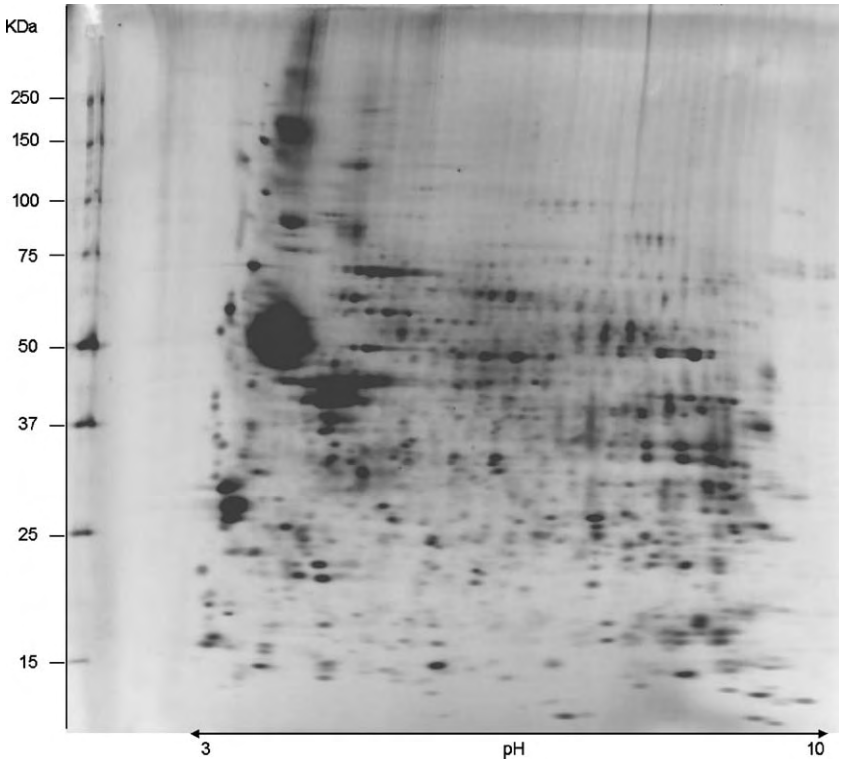

Figure 1 An example of rat hippocampus map. Proteins were focussed on 3—10 non-linear pH gradient strips and separated on 8—18.5% T polyacrylamide gels.

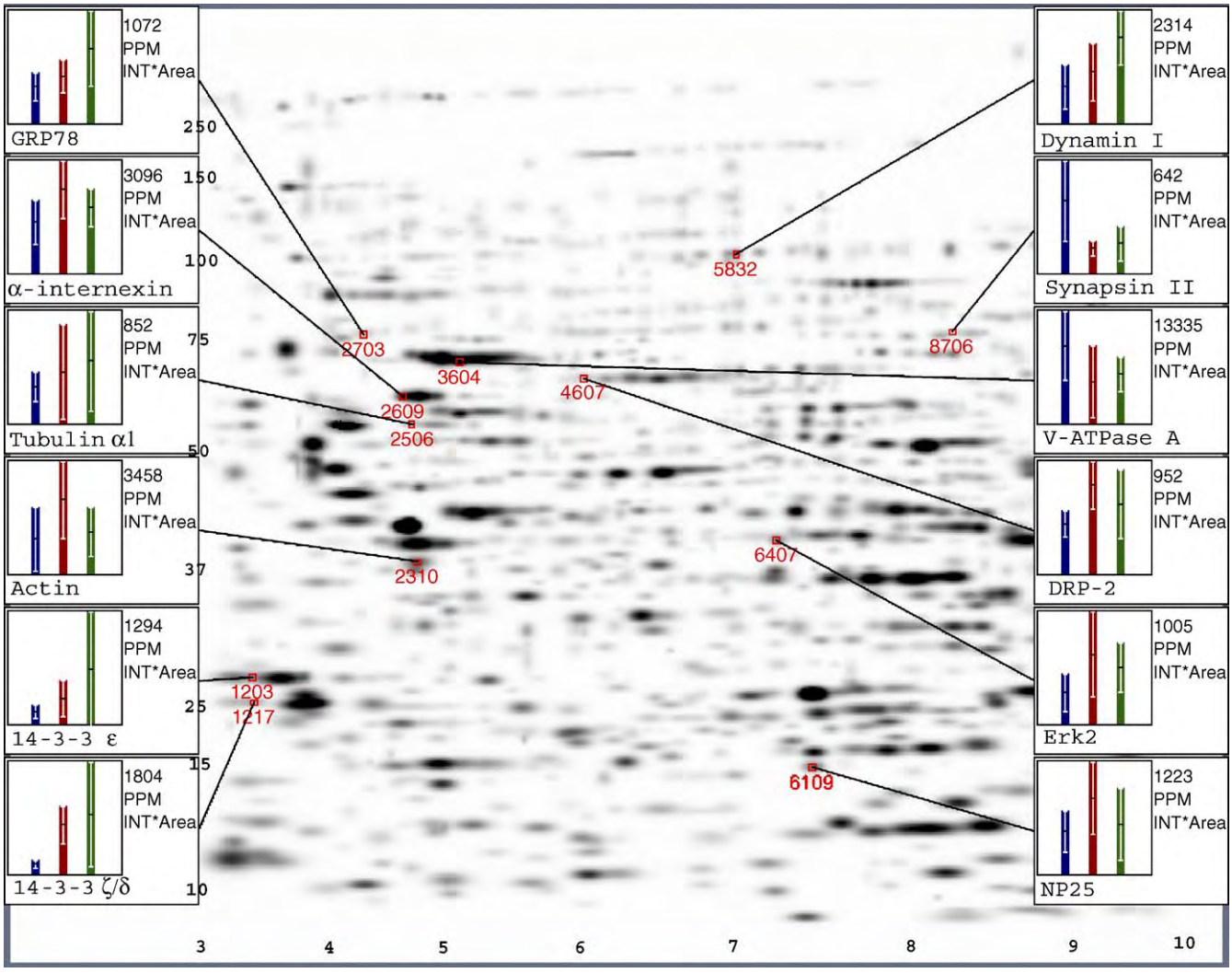

Figure 2 An example of rat frontal cortex map showing spots that displayed modified levels after pharmacological treatments. SSP (in red) is the identification number of the spot as reported in Tables 6 and 7. Bars show the average spot quantity (black dash) and standard deviation (vertical white lines). Blue bars: samples treated with vehicle, red bars: samples treated with fluoxetine, green bars: samples treated with GR205171. GRP78: 78 kDa glucose-regulated protein, V-ATPase A: vacuolar ATP synthase catalytic subunit A, DPR2: dihydropyriminidase-related protein 2, Erk2: mitogen-activated protein kinase 1, NP25: neuronal protein NP25. Numbers on the ordinate and abscissa axes correspond to approximate molecular mass and p*I* values, respectively.

acetonitrile). The extraction solutions were dried, redissolved in 10 μl 0.1% trifluoroacetic acid in water and purified by using ZIP-TIP C18 (Millipore Bedford, MA, USA). The purified peptides solution was added with matrix solution (10 mg/ml α-cyano-4-hydroxycinnamic acid in acetonitrile/ethanol, 1:1, v/v) and was analysed using a TofSpec 2E MALDI-TOF instrument (Micromass, Manchester, UK), operated in reflectron mode, with an accelerating voltage of 20 kV. Mascot (Matrix Science, UK), Profound (Version 4.10.5) and

Table 1 Quality data of the analysed matchsets: for each brain region and for the four matchsets, the number of detected spots per gel and the number of matched spots across gels are reported (average values ± standard deviations)

Brain region	Matchset	No. of detected spots per gel (average ± S.D.)	No. of matched spots across gels (average ± S.D.)	No. of matched spots in all gels	Mean coefficient of variation in the matchset
Hippocampus	Fluoxetine/GR205171/vehicle	644.5 ± 39.94	563.83 ± 38.18	246	38.4
	DMP696/vehicle	645.2 ± 60.80	534.87 ± 18.88	319	31.13
Cortex	Fluoxetine/GR205171/vehicle	819.79 ± 85.11	771.00 ± 72.92	317	46.45
	DMP696/vehicle	846.56 ± 41.40	788.56 ± 39.70	499	42.7

The number of matched spots in all gels of each matchset is also reported. Gels in a matchset were divided in different groups according to the treatment and the mean variation coefficient of spots across groups is reported for each matchset (standard deviation of spot quantity/average quantity × 100).

Table 2 Results obtained from the statistical analyses on gels

Brain region	Comparison set	TT ($P<0.05$)	PLS	Common to TT and PLS
Hippocampus	Fluoxetine vs. vehicle	6	21	2
	GR205171 vs. vehicle	8	19	4
	DMP696 vs. vehicle	98	16	7
Cortex	Fluoxetine vs. vehicle	42	40	17
	GR205171 vs. vehicle	33	14	8
	DMP696 vs. vehicle	81	14	9

Different comparison sets were investigated, comparing each treatment with its own vehicle. Two statistical tests were applied: Student's *t*-test (TT) and partial least squares test (PLS). The number of spots that showed a statistically significant modulation in expression level is reported for both tests. Modulated spots shared by both the tests are also indicated.

ProteinProbe (Version 3.4, BioLynx, Micromass, UK) softwares were used to search SwissProt, trEMBL and NCBI databases with Mammalia as taxonomic category to allow protein identification. Scores were reported as an indication of the probability of the identification. In ProteinProbe, the probability that a candidate in a database search is the protein being analysed is calculated in terms of Mowse Score. The protein can be considered identified when the Mowse scoring of the match is at least 10^5 or the score of the first candidate differs of three orders of magnitude from those of the others in the list given by the program. Mascot reports scores as $-10*\log_{10}(P)$, where P is the absolute probability that a match is random. Scores greater than 74 are significant ($p<0.05$), but given the power of Mascot, algorithm scores of 61 (or higher) indicate that there is a good likelihood of identification. ProFound calculates the probability in terms of *Z*-score, as an indicator of the quality of the search result. A *Z*-score of 1.65 for a search means that there are about 5% of random matches that could yield higher *Z*-scores than this search. The protein is likely to be identified when the *Z*-score of the match is 1.65 (or higher).

2.7. Western blot

Maps containing 700 μg of protein extract in 2D solubilising solution (see Section 2.3) were prepared from rat frontal cortex according to the method already described in Section 2.4. Gels were transferred on polyvinylidene fluoride membranes (Amersham Biosciences) for 3 h at 400 mA in 12.5 mM Tris, 96 mM glycine, 20% methanol, pH 8.2. After blocking non-specific binding in 5% non-fat dried milk in 0.1% Tween-20, Tris-buffered saline for 1 h at room temperature, membranes were incubated with 1:1000 anti-actin monoclonal antibody (Sigma-Aldrich) overnight at 4 °C under agitation. After extensive washing in 0.1% Tween-20 Tris-buffered saline, membranes were incubated with anti-mouse IgG coupled with horseradish peroxidase (Amersham Biosciences, 1:2000, 1 h at room temperature) and washes were repeated. Immunocomplexes were revealed by the ECL chemiluminescence method (Amersham Biosciences) following manufacturer's instructions.

3. Results

In this experimental scheme, rats were treated for 21 days with fluoxetine, a selective serotonin reuptake inhibitor, or with GR205171, a NK1 receptor antagonist, or with DMP696, a CRF receptor 1 antagonist. Control groups received the administration vehicle, which was water for fluoxetine and GR205171 and 0.5% methocel, 1% dimethyl-sulphoxide for DMP696. Protein extracts were prepared from hippocampus or frontal cortex and protein separation

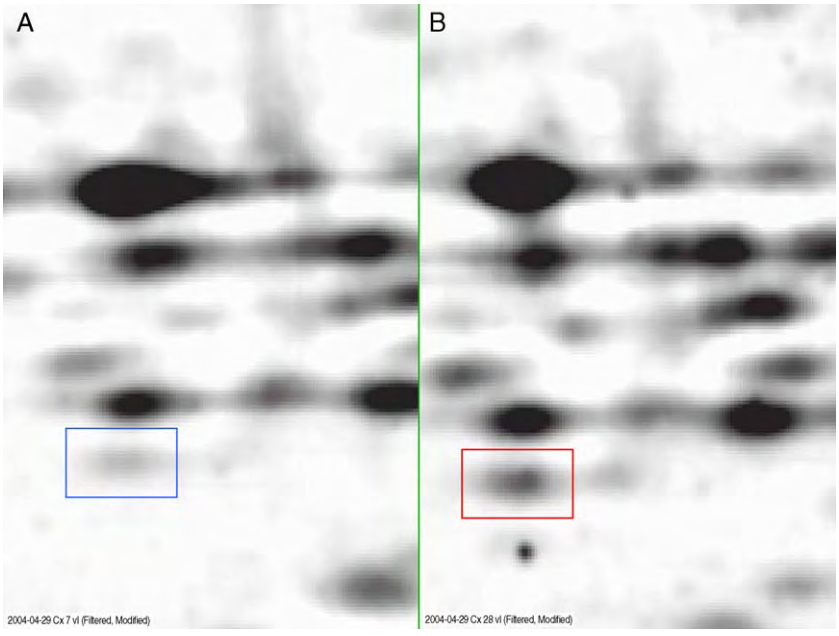

Figure 3 An example of spot quantity changes in rat cortex (synapsin II). (A) Vehicle treatment and (B) treatment with fluoxetine.

Table 3 Identification of six proteins displaying significantly different levels in rat hippocampus after exposure to chronic fluoxetine treatment

SSP	Mr (Da)	p*I*	MOWSE-score	MASCOT	*Z*-score	Protein name	Accession number	% Sequence coverage	No. of peptides	Regulation by drug treatment	%	TT	PLS
Cytoskeleton/neurofilaments													
2407	41,737	5.4	8,190,000	101	2.11	Actin, cytoplasmic 1	P60711	29	11	Down	50	No	Yes
2504	41,737	5.4	7.95E+08	76	2.03	Actin, cytoplasmic 1	P60711	28	10	Down	39	No	Yes
Signaling and cell cycle regulation													
1205	23,407	5.1	2.26E+08	157	2.39	Rho GDP-dissociation inhibitor 1	Q99PT1	54	15	Up	54	No	Yes
1202	27,754	4.8	2.19E+11	129	2.04	14-3-3 protein zeta/delta	P63102	48	17	Down	36	Yes	No
Carbohydrate transport/metabolism													
3311	36,481	6	2060	63	1.65	L-Lactate dehydrogenase B chain	P42123	14	7	Down	55	No	Yes
Nucleotide transport/metabolism													
8007	8123	9.6	13,600,000	138	1.78	ATP synthase e chain, mitochondrial	P29419	89	8	Up	45	Yes	No

SSP is the identification number of the selected spot assigned by the image analysis software in the specific matchset. Theoretical values of molecular mass (Mr) and isoelectric point (p*I*) are shown in columns 2 and 3. Probability scores of the identification are taken from the three database search softwares (see Section 2.6 for further explanations about their meaning). SwissProt protein names and accession numbers are provided in columns 7 and 8, respectively. The number of protein peptides that match the theoretical peptides from the database entry and the sequence coverage percentage are indicated in columns 10 and 9, respectively. The regulation refers to the drug-treated group as compared to the control group and the percentage of variation is shown in column 12. Statistical significance in TT or in PLS test are reported in columns 13 and 14, respectively.

Table 4 Identification of five proteins displaying significantly different levels in rat hippocampus after exposure to chronic GR205171 treatment

SSP	Mr (Da)	p*I*	MOWSE-score	MASCOT	*Z*-score	Protein name	Accession number	% Sequence coverage	No. of peptides	Regulation by drug treatment	%	TT	PLS
Cytoskeleton/neurofilaments													
2504	41,737	5.4	7.95E+08	76	2.03	Actin, cytoplasmic 1	P60711	28	10	Down	35	Yes	Yes
2705	49,943	5.4	1.79E+12	196	2.31	Glial fibrillary acidic protein, astrocyte	P47819	52	24	Up	63	No	Yes
Nucleotide transport/metabolism													
3002	18,632	6.5	4.49E+07	83	2.12	ATP synthase D chain, mitochondrial	P31399	51	9	Up	100	No	Yes
Energy production/metabolism													
8307	35,656	9.1	1.56E+06	116	2.37	Malate dehydrogenase, mitochondrial	P04636	42	13	Down	56	No	Yes
6601	57,650	6.7	1.49E+05	75	2.33	Pyruvate kinase	P11980	18	12	Up	54	Yes	No

Column labels are the same as in Table 3.

Table 5 Identification of 21 proteins displaying significantly different levels in rat hippocampus after exposure to chronic DMP696 treatment

SSP	Mr (Da)	p*I*	MOWSE-score	MASCOT	*Z*-score	Protein name	Accession number	% Sequence coverage	No. of peptides	Regulation by drug treatment	%	TT	PLS
Cytoskeleton/neurofilaments													
2401	41,737	5.4	2.76E+07	111	2.06	Actin, cytoplasmic 1	P60711	33	12	Up	68	Yes	No
2510	41,737	5.4	3.41E+08	105	2.20	Actin, cytoplasmic 1	P60711	34	14	Up	183	Yes	Yes
1702	49,963	4.8	6.43E+05	84	2.22	Tubulin beta chain	P04691	19	11	Up	58	Yes	No
Synaptic vesicles/ neurotransmitter transport													
1113	15,394	5.1	6.68E+04	71	1.65	Complexin 2	P84087	52	8	Up	55	Yes	No
Axonogenesis/membrane structures													
4703	62,278	6.3	7.95E+09	88	2.39	Dihydropyrimidinase-related protein-2	P47942	29	17	Down	40	Yes	No
Protein processing (folding, transport, metabolism)													
604	47,996	4.3	3.65E+04	93	2.30	Calreticulin	P18418	19	12	Up	82	Yes	No
9008	50,114	9.3	8.00E+03	83	1.80	Elongation factor 1-alpha 1	P62630	18	9	Down	40	No	Yes
8011	10,770	9.2	6.41E+06	165	2.41	10 kDa heat shock protein, mitochondrial	P26772	72	11	Up	63	Yes	No
1812*	84,657	5.0	5.97E+13	217	1.65	Heat shock protein HSP 90-alpha	P07901	32	24	Down	36	Yes	No
	83,194	5.0	4.91E+15			Heat shock protein HSP 90-beta	P11499	31	24				
1206	24,782	5.2	3.90E+06	102	1.74	Ubiquitin carboxyl-terminal hydrolase isozyme L1	Q00981	52	10	Up	94	Yes	No
Signaling and cell cycle regulation													
6	19,157	4.6	3.23E+03	60	1.67	Calcineurin B subunit isoform 1	P63100	28	5	Up	70	Yes	No
1202	27,754	4.8	2.19E+11	129	2.04	14-3-3 protein zeta/delta	P63102	48	17	Up	62	Yes	No
General metabolism/stimulus response													
4105	15,780	6.3	1.36E+04	74	1.65	Superoxide dismutase [Cu—Zn]	P07632	34	6	Up	75	Yes	No
Carbohydrate transport/metabolism													
3712	58,764	5.9	1.30E+06	93	2.30	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex	P08461	21	12	Down	34	Yes	No
1504	47,009	5.1	2.22E+06	115	2.40	Gamma enolase	P07323	24	12	Up	66	Yes	No
8401	39,221	8.5	8.30E+05	62	1.65	Fructose-bisphosphate aldolase A	P05065	41	11	Up	51	Yes	No
3404	36,481	6.0	7.66E+05	109	2.31	L-Lactate dehydrogenase B chain	P42123	32	12	Down	62	Yes	Yes
9302	35,656	9.1	7.33E+14	168	2.37	Malate dehydrogenase, mitochondrial	P04636	56	20	Up	60	Yes	No
6608	57,650	7.0	5.14E+10	99	2.16	Pyruvate kinase	P11980	37	19	Up	80	Yes	Yes
Unknown													
1401	24,936	4.8	1.05E+04	78	2.35	Protein C20orf178 homolog	Q9D8B3	34	9	Down	46	Yes	No

Column labels are the same as in Table 3.

Table 6 Identification of 35 proteins displaying significantly different levels in rat frontal cortex after exposure to chronic fluoxetine treatment

SSP	MW (Da)	p*I*	MOWSE-score	MASCOT	*Z*-score	Protein name	Accession number	% Sequence coverage	No. of peptides	Regulation by drug treatment	%	TT	PLS
Cytoskeleton/neurofilaments													
2310	41,710	5.3	2.05E+06	88	1.87	Actin, cytoplasmic 1	P60711	29	9	Up	131	Yes	Yes
2606*	56,115	5.3	2.69E+06	181	1.65	Alpha-internexin	P23565	26	12	Up	66	No	Yes
	49,910	5.0	2.29E+07			Tubulin alpha-2 chain	P05213	27	10				
2609	56,115	5.3	2.29E+10	152	2.39	Alpha-internexin	P23565	35	19	Up	64	Yes	No
3606*	56,115	5.3	6.56E+04	126	1.66	Alpha-internexin	P23565	19	8	Up	74	No	Yes
	49,943	5.4	6.11E+04			Glial fibrillary acidic protein, astrocyte	P47819	19	8				
1211*	35,613	4.9	1.18E+08	234	2.37	Annexin A5	P14668	38	14	Down	21	Yes	No
		5.0	*2.82E+06*			*Swiprosin 1*	*Q9D8Y0*	*40*	*10*				
6109	24,712	7.0	8.45E+05	78	1.65	Neuronal protein NP25	P37805	42	9	Up	77	Yes	Yes
4005	14,861	6.8	2.24E+03	105	2.34	Profilin-2	Q9EPC6	52	11	Up	20	Yes	No
2011	17,133	5.8	2.24E+03	62	1.65	Stathmin	P13668	42	8	Down	49	No	Yes
2506	49,910	5.0	9.68E+09	90	1.98	Tubulin alpha-1 chain	P68370	36	14	Up	42	No	Yes
Synaptic vesicles/ neurotransmitter transport													
8706	63,457	8.9	7.82E+15	98	2.40	Synapsin-2	Q63537	43	19	Down	65	Yes	Yes
Axonogenesis/membrane structures													
4607	62,278	6.3	1.34E+08	113	2.19	Dihydropyrimidinase-related protein-2	P47942	22	12	Up	76	Yes	No
Protein processing (folding, transport, metabolism)													
1502	47,996	4.3	1.08E+04	93	2.32	Calreticulin	P18418	16	9	Down	31	Yes	No
3619	56,623	6.1	3.46E+06	77	1.65	Protein disulfide-isomerase A3	P11598	22	12	Up	47	No	Yes
3602*	68,268	5.8	1.71E+12	160	2.32	*Vacuolar ATP synthase catalytic subunit A*	*P50516*	*34*	*18*	Up	73	Yes	No
	70,871	5.4	9.24E+10			Heat shock cognate 71 kDa protein	P63018	32	18				
1403	32,698	4.8	7.04E+03	72	1.85	40s ribosomal protein SA	P38983	32	9	Up	43	No	Yes

Signaling and cell cycle regulation												
3310*	37,307	5.6	4.00E+07	77	1.93	Guanine nucleotide-binding protein G(I)/G(S)/G(T) beta subunit 2	P54313	31	11	Down	39	No	Yes
	38,848	*6.2*				*Pyruvate dehydrogenase E1 component beta subunit, mitochondrial*	*P49432*	*31*	*10*				
1210	27,923	4.8	1.05E+12	118	2.04	14-3-3 protein zeta/delta	P63102	61	21	Down	30	Yes	No
1217	27,923	4.8	6.28E+05	105	2.29	14-3-3 protein zeta/delta	P63102	36	13	Up	400	Yes	No
1218	29,155	4.6	3.43E+08	69	2.16	14-3-3 protein epsilon	P62260	47	13	Down	47	Yes	No
6407	41,276	6.9	1.17E+05	72	1.94	Mitogen-activated protein kinase 1	P63086	23	10	Up	122	Yes	Yes
Amino acid transport/metabolism													
7411	46,168	6.3	5.36E+04	71	1.95	Aspartate aminotransferase, cytoplasmic	P13221	24	10	Up	28	No	Yes
Carbohydrate transport/metabolism													
4405	46,245	5.8	9.23E+05	87	1.91	Succinyl-CoA ligase [ADP-forming] beta-chain	Q9Z2I9	23	12	Up	72	Yes	No
Nucleotide transport/metabolism													
7002	17,283	7.6	1.05E+05	85	1.91	Nucleoside diphosphate kinase B	P19804	48	8	Up	40	Yes	No
3604	68,268	5.8	1.07E+15	123	2.36	Vacuolar ATP synthase catalytic subunit A	P50516	39	24	Down	46	Yes	Yes
3608	68,268	5.8	1.09+20	155	2.41	Vacuolar ATP synthase catalytic subunit A	P50516	44	28	Up	160	Yes	No
1310	40,301	4.9	4.33E+04	98	2.24	Vacuolar ATP synthase subunit d	P51863	19	8	Down	61	Yes	No
Energy production/metabolism													
6714	85,358	8.3	4.23E+04	91	2.03	Aconitate hydratase, mitochondrial	P20004	12	7	Up	57	No	Yes
General metabolism/stimulus response													
6106	23,308	7.5	4.80E+08	75	1.65	Glutathione S-transferase P	P04906	68	10	Up	71	Yes	No
7106	25,550	7.5	2.50E+08	170	2.23	Glutathione S-transferase Yb-3	P08009	55	13	Up	59	Yes	No

Column labels are the same as in Table 3.

Table 7 Identification of 14 proteins displaying significantly different levels in rat frontal cortex after exposure to chronic GR205171 treatment

SSP	MW (Da)	pI	MOWSE-score	MASCOT	Z-score	Protein name	Accession number	% Sequence coverage	No. of peptides	Regulation by drug treatment	%	TT	PLS
Cytoskeleton/neurofilaments													
5832	95,867	6	1.11E+10	144	2.41	Dynamin-1	P21575	20	21	Up	133	Yes	Yes
Synaptic vesicles/neurotransmitter transport													
8706	63,457	9	7.82E+15	98	2.4	Synapsin-2	Q63537	43	19	Down	59	Yes	Yes
Axonogenesis/membrane structures													
6617*	62,278	6	3.00E+13	208	1.72	Dihydropyrimidinase-related protein-2	P47942	41	19	Up	87	Yes	No
	68,692	*6*	*2.27E+11*			*Syntaxin binding protein 1*	*P61765*	*32*	*18*				
Protein processing (folding, transport, metabolism)													
4603	56,623	6	1.35E+08	80	2.34	Protein disulfide-isomerase A3	P11598	27	15	Down	39	Yes	Yes
2410	52,769	6	5.92E+07	66	2.33	Ubiquinol-cytochrome-*c* reductase complex core protein I	Q9CZ13	29	14	Up	53	Yes	No
2703	72,347	5	8.25E+11	171	2.42	78 kDa glucose-regulated protein	P06761	28	18	Up	101	Yes	No
Signaling and cell cycle regulation													
2313	39,949	6	2.35E+10	64	2.3	Guanine nucleotide-binding protein G(o), alpha subunit 2	P30033	40	15	Up	62	Yes	No
4409	50,537	6	8.92E+05	69	1.65	Rab GDP dissociation inhibitor beta-2	Q61598	25	10	Up	54	Yes	No
1203	29,155	5	8.90E+05	93	1.89	14-3-3 protein epsilon	P62260	36	10	Up	334	Yes	No
1218	29,155	5	3.43E+08	69	2.16	14-3-3 protein epsilon	P62260	47	13	Down	53	Yes	No
Carbohydrate transport/metabolism													
5312	35,643	7	2.24E+03	60	2.33	Aldose reductase	P07943	15	7	Up	157	Yes	Yes
Energy production/metabolism													
7706	85,358	8	6.05E+05	120	2.3	Aconitate hydratase, mitochondrial	P20004	14	10	Up	33	Yes	No
General metabolism													
6106	23,308	8	4.80E+08	75	1.65	Glutathione *S*-transferase P	P04906	68	10	Up	64	Yes	No

Column labels are the same as in Table 3.

was achieved on 2D gels, based on the isoelectric point and molecular mass. Proteins separated on the maps were revealed by staining with Sypro Ruby fluorescent dye and image analysis was carried out on the acquired image. Maps belonging to the same run were compared to each other by creating matchsets including 24 or 16 maps each. In each matchset, a master image was created, which contained most of the spots. Corresponding spots in all gels were matched and attributed an identification number, which was specific of the matchset. Sample gels for hippocampus or frontal cortex proteins are shown in Figs. 1 and 2, respectively. Four matchsets were created, grouping antidepressant treatments with their own vehicle, for both brain regions. Information about the quality of the obtained matchsets are indicated in Table 1, where the number of detected spots per gel and the number of matched spots across gels are reported for each matchset. Gels in a matchset were divided in different groups according to the treatment and the variation coefficient of spots across groups was calculated (standard deviation of spot quantity / average quantity × 100), as reported in Table 1. Treatment groups were compared with their vehicles and statistical analyses were carried out on log transformed values since previous analyses suggested that a normal distribution was achieved after the transformation. A Student's *t*-test and a partial least squares analysis were performed for each comparison set in order to evaluate the results obtained with univariate and multivariate methods. Results from the statistical analyses are reported in Table 2, where the number of modulated spots for each treatment is indicated, according to the used statistical test. As shown in Table 2, some spots were common between the two statistical approaches in the same comparison set. For the drug treatments run in the same matchset (i.e. fluoxetine and GR205171), we could identify spots modulated in common: two spots were found to be modulated by both drugs in hippocampus and 11 spots in cortex, considering either spots detected by TT and PLS. An example of modulated spot is shown in Fig. 3. Spots displaying statistically significant variations in expression levels were subjected to peptide fingerprinting mass spectrometry to identify them. Spots showing very low changes (lower than 15%) were discarded because they were not considered sufficiently meaningful from a biological point of view, although they were statistically significant. Successful identifications are reported in Tables 3–8. An example of the location of identified actin spots in a frontal cortex 2D map is shown in Fig. 4.

4. Discussion

This study was undertaken to gain insight into the molecular modifications induced by chronic antidepressant treatment and with two compounds belonging to new classes of potential antidepressants. The global pattern of protein expression was compared after chronic treatments by means of 2D electrophoresis followed by image analysis and statistical evaluation. Proteomic approaches produce large datasets, which offer challenges to the effort of extracting as much biological information as possible, and unanimous agreement has not been reached about, which is the best method to use for the statistical analysis. In particular, both univariate and multivariate data analysis methods have been suggested, with both of them showing strengths and weaknesses (Karp et al., 2005; Klenø et al., 2004). In the univariate approach, each variable is analysed independently of the others, thus identifying significant changes in specific spots. Multivariate methods look at the global picture and they are more apt at discovering changes in the general pattern of protein expression. The two approaches are thus complementary and, since they are based on different methods, changes detected in common are likely to be truly significant (Karp et al., 2005). In this work, we adopted the strategy of analysing data with both univariate and multivariate methods, that is the Student's *t*-test (TT) and the partial least squares analysis (PLS). Although both methods are prone to give a number of false positive results, the changes that were detected by both methods were considered as more relevant. Our focussing on overlapping modifications does not imply that changes detected by TT or PLS alone may not bear a biological relevance; it only means that there is a lower confidence that those modifications were truly induced by the treatment.

A general observation is that a low number of changes was induced in brain protein levels by the treatment, as already reported by other studies (Guest et al., 2004; Kawaja et al., 2004). This finding is not surprising, since the experimental procedures involved treatment of naïve animals with a pharmacological dosage that does not induce overt phenotypic changes besides changes in behaviour detectable only through specific tests.

An objective of this study was the comparison of changes induced by the different compounds, in order to assess whether new classes of putative antidepressants could induce similar or unique changes when compared to fluoxetine. We detected that a protein modulation shared by all three compounds involves changes in levels of actin isoforms, also shown in Fig. 4. Recent studies on the pathophysiology of mood disorders and the mechanism of action of antidepressive agents suggested the involvement of alterations in neural plasticity (Manji et al., 2001; Duman, 2002; D'Sa and Duman, 2002). Since the regulation of actin cytoskeleton plays a major role in neuronal morphogenesis and in synaptic plasticity (Luo, 2002; Matus, 2000), it is plausible that the changes we detected can be explained in this framework. The modification of actin isoform levels are possibly involved in the synaptic remodelling induced by chronic treatment with antidepressants (Guest et al., 2004). However, as the proteomic analysis has been run on tissue homogenates and not on subcellular fractions, it is not possible to establish if the changes detected are strictly arising from modulation of synaptic actin or to other plasticity events impacting on actin levels. Since several antidepressants and electroconvulsive treatment are reported to be able to alter neuroplasticity (Popoli et al., 2002; Fuchs et al., 2004) and to induce the expression of genes involved in neural plasticity (Palotàs et al., 2004), it is possible that the regulation of this mechanism is related to the efficacy of an antidepressive agent. This hypothesis is consistent with the observation that both fluoxetine and GR205171 reduced synapsin II levels. Synapsins are neuronal phosphoproteins associated with the surface of synaptic

Table 8 Identification of 45 proteins displaying significantly different levels in rat frontal cortex after exposure to chronic DMP696 treatment

SSP	pI	MOWSE-score	MASCOT	Z-score	Protein name	Accession number	% Sequence coverage	No. of peptides	Regulation by drug treatment	%	TT	PLS
Cytoskeleton/neurofilaments												
2319	4.8	1.49E+14	134	2.4	Actin, cytoplasmic 1	P60711	56	19	Down	20	Yes	No
1307	4.8	2.05E+06	93	1.7	Actin, cytoplasmic 1	P60711	29	9	Down	32	Yes	No
2512	6	5.60E+16	223	2.32	Alpha internexin	P23565	56	32	Up	50	Yes	No
6801	6.7	1.16E+08	131	2.38	Dynamin-1	P21575	16	21	Down	37	Yes	No
5811	6.7	7.34E+09	95	2.34	Dynamin-1	P21575	18	24	Down	35	Yes	No
5701	5.3	1.50E+06	102	1.65	Ezrin	P26040	13	13	Down	34	Yes	No
5705	6	1.25E+10	105	2.4	Ezrin	P26040	31	23	Down	33	Yes	No
1513	5	2.47E+16	146	2.39	Tubulin beta chain	P04691	43	24	Down	20	Yes	No
2501	5.3	1.14E+08	89	1.89	Tubulin beta chain	P04691	31	13	Down	20	Yes	No
2508	5.3	1.70E+33	187	2.43	Tubulin alpha-2 chain	P05213	59	35	Down	32	Yes	No
Synaptic vesicles/neurotransmitter transport												
1804	4.6	5.33E+08	91	2.39	Amphiphysin	O08838	22	13	Up	27	Yes	No
1204	4.6	5.67E+03	89	2.02	Clathrin light chain B	P08082	24	8	Up	109	Yes	No
6611	7	1.03E+09	116	2.32	Syntaxin binding protein 1	P61765	27	19	Down	33	Yes	No
6614	7	3.11E+08	88	2.26	Syntaxin binding protein 1	P61765	26	17	Up	39	Yes	No
7711	8.9	3.35E+08	80	2.41	Synapsin-2	Q63537	29	11	Up	66	Yes	No
7504*	*7.6*	*4.08E+08*	147	1.65	*Pyruvate kinase*	P14618	26	12	Up	67	Yes	No
	8.9	1.52E+09			Synapsin-2	Q63537	27	10				
Axonogenesis/membrane structures												
5607	6.3	2.42E+24	201	2.39	Dihydropyrimidinase-related protein-2	P47942	57	31	Down	33	Yes	No
5711	6.3	1.13E+08	74	2.2	Dihydropyrimidinase-related protein-2	P47942	27	12	Down	54	Yes	Yes
4002	11.4	2.48E+03	82	1.74	Myelin basic protein S	P02688	44	6	Down	32	Yes	No
Protein processing (folding, transport, metabolism)												
1808	4.8	1.89E+06	84	2.36	Endoplasmin	P08113	16	17	Up	37	Yes	No
4610*	6.9	4.71E+05	69	2.4	Heat shock protein 75 kDa	Q9CQN1	15	10	Down	43	No	Yes
	8.6	7.04E+03	71		*40s ribosomal protein SA*	P38983	32	9				
8001	4.8	4.98E+07	81	1.65	Peptidyl-prolyl *cis—trans* isomerase A	P10111	51	11	Up	33	Yes	No
6706	6.6	5.97E+05	68	2	Vesicle-fusing ATPase	P46460	17	12	Down	45	Yes	No

Signaling and cell cycle regulation												
6309	6.9	6.75E+07	119	2.36	Mitogen-activated protein kinase 1	P63086	34	13	Up	89	Yes	No
1108	4.8	2.62E+05	94	1.89	Translationally controlled tumor protein	P63029	40	8	Down	33	Yes	No
1109	4.8	2.85E+11	123	1.99	14-3-3 protein zeta/delta	P63102	60	20	Up	115	Yes	No
Amino acid transport/metabolism												
6306	7.3	2.74E+07	86	2.06	Glutamine synthetase	P09606	28	12	Down	19	Yes	No
7502	8.3	1.31E+15	140	2.35	Glutamate dehydrogenase 1	P10860	45	22	Down	26	Yes	No
7305*	6.8	7.06E+15	188	2.21	*Fructose-bisphosphate aldolase C*	P09117	53	15	Down	28	Yes	No
	6.3	2.11E+02			Aspartate aminotransferase, cytoplasmic	P13221	29	13				
Carbohydrate transport/metabolism												
5402	5.8	2.12E+14	111	2.24	Alpha enolase	P04764	45	19	Down	20	Yes	No
4409	6.5	8.18E+08	127	2.27	Alpha enolase	P04764	37	13	Up	55	Yes	No
7314	8.5	1.22E+08	83	Low score	Fructose-bisphosphate aldolase A	P05065	41	11	Up	200	Yes	No
7306	8.6	9.56E+13	84	2.25	Glyceraldehyde 3-phosphate dehydrogenase	P04797	51	19	Up	52	Yes	No
Nucleotide transport/metabolism												
8504	9.4	2.65E+20	332	2.36	ATP synthase alpha chain	P15999	61	33	Up	61	Yes	No
8519	9.4	6.63E+05	85	1.74	ATP synthase alpha chain	P15999	19	11	Down	57	Yes	No
7510	9.4	5.49E+13	170	2.33	ATP synthase alpha chain	P15999	49	24	Up	118	Yes	Yes
1308	4.9	2.49E+08	133	2.33	Vacuolar ATP synthase subunit d	P51863	37	14	Up	21	Yes	No
Energy production/metabolism												
7706	6.9	1.36E+10	86	2.32	Aconitate hydratase, mitochondrial	P20004	20	15	Up	39	Yes	No
3411	5.5	2.23E+13	98	2.38	Creatine kinase, B chain	P07335	49	16	Up	85	Yes	No
4415	8.9	Low score	Low score	Low score	NADH-ubiquinone oxidoreductase 49 kDa subunit	Q91WD5	Low score	Low score	Down	43	Yes	No
7514	8.3	2.63E+09	148	2.37	Pyruvate kinase	P14618	33	16	Down	47	Yes	Yes

Column labels are the same as in Table 3.

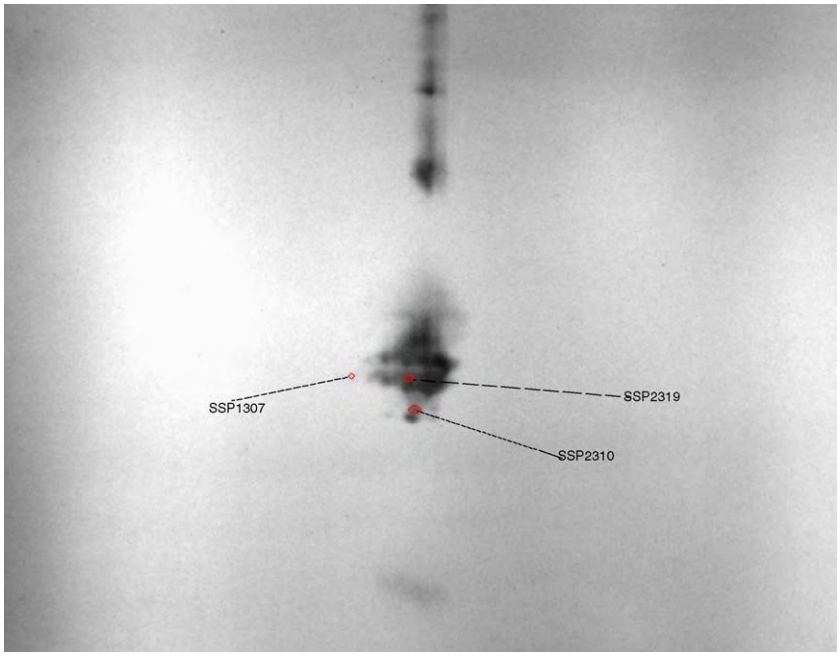

Figure 4 An example of 2D gel prepared from rat frontal cortex protein extract blotted with anti-actin antibody. SSP is the identification number of the spot as reported in Tables 6—8.

vesicles. They are involved in the control of the availability of synaptic vesicles for exocytosis and also in the regulation of neurotransmitter release (Greengard et al., 1993; Hilfiker et al., 1999). Synapsin II as well as synapsin I are involved in mediating the brain-derived neurotrophic factor-mediated modulation of glutamate release (Jovanovic et al., 2000). Thus, the detected reduction of synapsin II levels would be able to induce a modulation of synaptic function that could be involved in mediating the antidepressant action of the compounds.

Among the changes in protein levels which were detected after chronic fluoxetine treatment, we identified an increase in ERK2 (also known as MAPK1), that was also found to be modulated by DMP696. ERK2 belongs to a family of parallel cascades of kinases activated by phosphorylation, which respond to extracellular stimuli by inducing cellular responses (Raman and Cobb, 2003; Pearson et al., 2001). In particular, ERK2 belongs to a pathway often activated by survival signals, among which growth factors, such as brain-derived neurotrophic factor (Segal and Greenberg, 1996). In brain, ERKs play a well-recognised role in mechanisms of synaptic plasticity, especially in the molecular bases of memory formation (Sweatt, 2004; Thomas and Huganir, 2004). Modifications in ERK levels or phosphorylation were already reported in rodent brain regions after antidepressant treatment (Tiraboshi et al., 2004; Valjent et al., 2004). Moreover, a reduction of ERK2 levels was discovered in the same brain regions in a model of depression in rats (Feng et al., 2003). These findings support the hypothesis theories that antidepressant action is associated with the ability to sustain neurotrophic and neurogenetic signalling (Reid and Stewart, 2001; Castrén, 2004). Another protein showing up-regulated levels after chronic fluoxetine treatment is NP25 (also known as Transgelin 3). NP25 is a neuronal protein of unknown function whose mRNA is expressed at high levels in rat hippocampus and cortex (Ren et al., 1994). Although the protein is poorly characterised, thus understanding its function is difficult, recently, its interaction with F-actin was described (Mori et al., 2004), thus suggesting its involvement in morphological remodelling.

Within the hypothesis of antidepressant acting as modulators of synaptic plasticity also has place the detected reduction of vacuolar ATP synthase. Among proteins involved in synaptic activity, vacuolar ATP synthase belongs to a family of enzymes responsible for the acidification of intracellular compartments, mediating endocytosis and intracellular transport (Nishi and Forgac, 2002). In brain, they are involved in the re-uptake of neurotransmitters into synaptic vesicles (Maycox et al., 1990).

Specific changes induced by GR205171 are the reduction of protein disulphide isomerase A, which contributes to the conformation of proteins, the increase of dynamin 1, a brain-specific GTPase involved in vesicular trafficking, and the increase of aldose reductase, an enzyme that catalyses the reduction of aldehydes to sugars.

Among changes detected after chronic DMP696 administration, we found a reduction of dihydropyrimidinase-related protein-2 (DRP-2) levels. DRP-2 is a substrate of Ca^{2+}/calmodulin kinase II in the post-synaptic density of adult rat brain, thus suggesting that the protein plays a role in synaptic transmission and plasticity (Yoshimura et al., 2002). DRP-2 is modified by oxidative damage in Alzheimer's disease brain (Castegna et al., 2002) and a reduced amount was revealed in brains of patients affected with major depressive disorder, bipolar disorder, schizophrenia and Down's syndrome (Johnston-Wilson et al., 2000; Lubec et al., 1999; Weitzdoerfer et al., 2001). In addition, the antidepressive agents venlafaxine or fluoxetine increase DRP-2 in rat hippocampus (Kawaja et al., 2004).

In agreement with Kawaja et al. (2004), we also detected an increase in DRP-2 after chronic fluoxetine treatment,

although it was detected only in the statistical analysis with TT. The reduction detected after GR205171 treatment apparently goes in the opposite direction, although caution should be used in interpreting this finding, because DRP-2 belongs to a group of proteins that often show changes after several kind of perturbations (Fountoulakis, 2004). The other changes concern proteins involved in energy metabolism. Among them, pyruvate kinase showed changes in opposite directions in the two brain areas examined, suggesting a modulation of post-translationally modified forms or a region-specific modulation. A modification of this enzyme after chronic paroxetine treatment was detected also at the mRNA levels by Landgrebe et al. (2002). It is noteworthy that in both regions the treatment with DMP696 induced a large number of modifications when analysed with TT, but only a few of them can also be revealed by PLS test, possibly because of the different method used for the analysis.

As already stated, the proteins pointed out by both method of data analysis can be related to the pharmacological treatment with a higher degree of confidence. It is worth considering that for some changes that were detected with only one method, a relationship with antidepressant treatment has already been reported in other studies. As an example, while we detected an increase in glyceraldehyde-3-phosphate dehydrogenase protein levels after treatment with DMP696, other studies reported that chronic treatment with antidepressants induced an increase in glyceraldehyde-3-phosphate dehydrogenase mRNA (Tohda et al., 1999; Drigues et al., 2003). The increase in Cu/Zn superoxide dismutase induced by DMP696 was also detected after chronic treatment with venlafaxine or amitriptyline (Xu et al., 2003). We also detected a modulation of the levels of G protein subunits after fluoxetine or GR205171 treatment, which is in keeping with the involvement detected for G proteins in the action of antidepressants (Galeotti et al., 2002; Donati and Rasenick, 2003).

In conclusion, the proteomic analysis of rat frontal cortex and hippocampus after chronic treatment with an SSRI antidepressant, a NK1 antagonist and a CRF receptor 1 antagonist revealed a specific pattern of protein modulation for each pharmacologically active compound. The common theme that was detected in changes induced by each treatment supports the hypothesis that efficacious antidepressants share the ability of modulating neural plasticity.

Acknowledgements

We thank Dr. Mahmoud Hamdan for kind advice in mass spectrometry experiments, Prof. Pier Giorgio Righetti, Dr. Mauro Corsi and Dr. Roberto Arban for helpful discussions, and Dr. David Wille' for statistical advice.

References

Adell, A., 2004. Antidepressant properties of substance P antagonists: relationship to monoaminergic mechanisms? Curr. Drug Targets CNS Neurol. Disord. 3, 113–121.

Alonso, R., Griebel, G., Pavone, G., Stemmelin, J., Le Fur, G., Soubrie, P., 2004. Blockade of CRF(1) or V(1b) receptors reverses stress-induced suppression of neurogenesis in a mouse model of depression. Mol. Psychiatry 224, 278–286.

Castegna, A., Aksenov, M., Thongboonkerd, V., Klein, J.K., Pierce, W.M., Booze, R., Markesbery, W.R., Butterfield, D.A., 2002. Proteomic identification of oxidatively modified proteins in Alzheimer's disease brain: Part II. Dihydropyrimidinase-related protein 2, α-enolase and heat-shock cognate 71. J. Neurochem. 82, 1524–1532.

Castrén, E., 2004. Neurotrophic effects of antidepressant drugs. Curr. Opin. Pharmacol. 4, 1–7.

Claes, S.J., 2004. Corticotropin-releasing hormone (CRH) in psychiatry: from stress to psychopathology. Ann. Med. 36, 50–61.

Donati, R., Rasenick, M.M., 2003. G protein signaling and the molecular basis of antidepressant action. Life Sci. 73, 1–17.

Drevets, W.C., Price Jr., J.L., Simpson, J.R., Todd, R.D., Reich, T., Vannier, M., Raichle, M.E., 1997. Subgenual prefrontal cortex abnormalities in mood disorders. Nature 386, 824–827.

Drigues, N., Poltyrev, T., Bejar, C., Weinstock, M., Youdim, M.B.H., 2003. cDNA gene expression profile of rat hippocampus after chronic treatment with antidepressant drugs. J. Neural Transm. 110, 1413–1436.

D'Sa, C., Duman, R.S., 2002. Antidepressants and neuroplasticity. Bipolar Disord. 4, 183–194.

Duman, R.S., 2002. Pathophysiology of depression: the concept of synaptic plasticity. Eur. Psychiatry 17 (Suppl 3), 306–310.

Feng, P., Guan, Z., Yang, X., Fang, J., 2003. Impairments of ERK signal transduction in the brain in a rat model of depression induced by neonatal exposure of clomipramine. Brain Res. 991, 195–205.

Fountoulakis, M., 2004. Application of proteomics technologies in the investigation of the brain. Mass Spectrom. Rev. 23, 231–258.

Fuchs, E., Czeh, B., Kole, M.H.P., Michaelis, T., Lucassen, P.J., 2004. Alterations of neuroplasticity of depression: the hippocampus and beyond. Eur. Neuropsychopharmacol. 14, S481–S490.

Galeotti, N., Bartolini, A., Ghelardini, C., 2002. Role of G1 proteins in the antidepressant-like effect of amitriptyline and clomipramine. Neuropsychopharmacology 27, 554–564.

Gardner, C.J., Armour, D.R., Beattie, D.T., Gale, J.D., Hawcock, A.B., Kilpatrick, G.J., Twissell, D.J., Ward, P., 1996. GR205171: a novel antagonist with high affinity for the tachykinin NK1 receptor, and potent broad-spectrum anti-emetic activity. Regul. Pept. 65, 45–53.

Gram, L., 1994. Fluoxetine. N. Engl. J. Med. 331, 1354–1361.

Greengard, P., Valtorta, F., Czernik, A.J., Benfenati, F., 1993. Synaptic vesicle phosphoproteins and regulation of synaptic function. Science 259, 780–785.

Guest, P.C., Knowles, M.R., Molon-Noblot, S., Salim, K., Smith, D., Murray, F., Laroque, P., Hunt, S.P., de Felipe, C., Rupniak, M.N., McAllister, G., 2004. Mechanism of action of the antidepressant fluoxetine and the substance P antagonist L-000760735 are associated with altered neurofilaments and synaptic remodeling. Brain Res. 1002, 1–10.

He, L., Gilligan, P.J., Zaczek, R., Fitzgerald, L.W., McElroy, J., Shen, H.-S.L., Saye, J.A., Kalin, N.H., Shelton, S., Christ, D., Trainor, G., Hartig, P., 2000. 4-(1,3-Dimethoxyprop-2-ylamino)-2,7-dimethyl-8-(2,4-dichlorophenyl)-pyrazolo[1,5-*a*]-1,3,5-triazine: a potent, orally bioavailable CRF_1 receptor antagonist. J. Med. Chem. 43, 449–456.

Herbert, B., Galvani, M., Hamdan, M., Olivieri, E., MacCarthy, J., Pedersen, S., Righetti, P.G., 2001. Reduction and alkylation of proteins in preparation of two-dimensional map analysis: why, when and how? Electrophoresis 22, 2046–2057.

Hilfiker, S., Pieribone, V.A., Czernik, A.J., Kao, H.-T., Augustine, G.J., Greengard, P., 1999. Synapsins as regulators of neurotransmitter release. Philos. Trans. R. Soc. Lond., B 354, 269–279.

Holsboer, F., 1999. The rationale for corticotropin-releasing hormone receptor (CRH-R) antagonists to treat depression and anxiety. J. Psychiatr. Res. 33, 181–214.

Holsboer, F., 2003. Corticotropin-releasing hormone modulators and depression. Curr. Opin. Investig. Drugs 4, 46–50.

Jovanovic, J.N., Czernik, A.J., Fienberg, A.A., Greengard, P., Sihra, T.S., 2000. Synapsins as mediators of BDNF-enhanced neurotransmitter release. Nat. Neurosci. 3, 323–329.

Johnston-Wilson, N.L., Sims, C.D., Hofman, J-P., Anderson, L., Shore, A.D., Torrey, E.F., Yolken, R.H., 2000. Disease-specific alterations in frontal cortex brain proteins in schizophrenia, bipolar disorder, and major depressive disorder. Mol. Psychiatry 5, 142–149.

Karp, N.A., Griffin, J.L., Lilley, K.S., 2005. Application of partial least squares discriminant analysis to two-dimensional difference gels studies in expression proteomics. Proteomics 5, 81–90.

Kawaja, X., Xu, J., Liang, J.-J., Barrett, J.E., 2004. Proteomic analysis of protein changes developing in rat hippocampus after chronic antidepressant treatment: implications for depressive disorders and future therapies. J. Neurosci. Res. 75, 451–460.

Kehne, J., De Lombaert, S., 2002. Non-peptidic CRF_1 receptor antagonists for the treatment of anxiety, depression and stress disorders. Curr. Drug Targets CNS Neurol. Disord. 1, 467–493.

Klenø, T.G., Leonardsen, L.R., Kjeldal, H.Ø., Laursen, S.M., Jensen, O.N., Baunsgaard, D., 2004. Mechanisms of hydrazine toxicity in rat liver investigated by proteomics and multivariate data analysis. Proteomics 4, 868–888.

Landgrebe, J., Welzl, G., Metz, T., van Gaalen, M.M., Ropers, H., Wurst, W., Holsboer, F., 2002. Molecular characterisation of antidepressant effects in the mouse brain using gene expression profiling. J. Psychiatr. Res. 36, 119–129.

Lubec, G., Nonaka, M., Krapfenbauer, K., Gratzer, M., Cairns, N., Fountoulakis, M., 1999. Expression of the dihydropyrimidinase-related protein 2 (DRP-2) in Down syndrome and Alzheimer's disease brain is downregulated at the mRNA and dysregulated at the protein level. J. Neural Transm., Suppl. 57, 161–177.

Luo, L., 2002. Actin cytoskeleton regulation in neuronal morphogenesis and structural plasticity. Annu. Rev. Cell. Dev. Biol. 18, 601–635.

Manji, H.K., Drevets, W.C., Charney, D.S., 2001. The cellular neurobiology of depression. Nat. Med. 7, 541–547.

Matus, A., 2000. Actin-base plasticity in dendritic spines. Science 290, 754–758.

Maycox, P.R., Hell, J.W., Jahn, R., 1990. Amino acid neurotransmission: spotlight on synaptic vesicles. Trends Neurosci. 13, 83–87.

Mori, K., Muto, Y., Kokuzawa, J., Yoshioka, T., Yoshimura, S., Iwama, T., Okano, Y., Sakai, N., 2004. Neuronal protein NP25 interacts with F-actin. Neurosci. Res. 48, 439–446.

Müller, M., Wurst, W., 2004. Getting closer to affective disorders: the role of CRH receptor systems. Trends Mol. Med. 10, 409–415.

Nemeroff, C.B., 2002. New directions in the development of antidepressants: the interface of neurobiology and psychiatry. Hum. Psychopharmacol. 17, S13–S16.

Nemeroff, C.B., Owens, M.J., 2002. Treatment of mood disorders. Nat. Neurosci. Suppl. 5, 1068–1070.

Nestler, E.J., 1998. Antidepressant treatments in the 21st century. Biol. Psychiatry 44, 526–533.

Nishi, T., Forgac, M., 2002. The vacuolar (H+)-ATPases—nature's most versatile proton pumps. Nat. Rev., Mol. Cell Biol. 3, 94–103.

Palotàs, M., Palotàs, A., Puskàs, L.G., Kitajka, K., Pàkàski, M., Janka, J., Penke, B., Kàlmàn, J., 2004. Gene expression profile analysis of the rat cortex following treatment with imipramine and citalopram. Int. J. Neuropsychopharmacol. 7, 1–13.

Pearson, G., Robinson, F., Beers Gibson, T., Xu, B.-E., Karandikar, M., Berman, K., Cobb, M.H., 2001. Mitogen-activated protein (MAP) kinase pathways: regulation and physiological functions. Endocr. Rev. 22, 153–183.

Piubelli, C., Cecconi, D., Astner, H., Caldara, F., Tessari, M., Carboni, L., Hamdan, M., Righetti, P.G., Domenici, E., 2005. Proteomic changes in rat serum, polymorphonuclear and mononuclear leukocytes after chronic nicotine administration. Proteomics 5, 1382–1394.

Popoli, M., Gennarelli, M., Racagni, G., 2002. Modulation of synaptic plasticity by stress and antidepressants. Bipolar Disord. 4, 166–182.

Posener, J.A., Wang, L., Price, J.L., Gado, M.H., Province, M.A., Miller, M.I., Babb, C.M., Csernansky, J.G., 2003. High-dimensional mapping of the hippocampus in depression. Am. J. Psychiatry 160, 83–89.

Raman, M., Cobb, M.H., 2003. MAP kinase modules: many roads home. Curr. Biol. 13, R886–R888.

Reid, I.C., Stewart, C.A., 2001. How antidepressants work. Br. J. Psychiatry 178, 299–303.

Ren, W.-Z., Ng, G.Y.K., Wang, R.-X., Wu, P.H., O'Dowd, B.F., Osmond, D.H., Goerge, S., Liew, C.-C., 1994. The identification of NP25: a novel protein that is differentially expressed by neuronal subpopulations. Mol. Brain Res. 22, 173–185.

Rupniak, N.M.J., 2002. Elucidating the antidepressant actions of the substance P (NK_1 receptor) antagonists. Curr. Opin. Investig. Drugs 3, 257–261.

Rupniak, N.M.J., Carlson, E.J., Shepheard, S., Bentley, G., Williams, A.R., Hill, A., Swain, C., Mills, S.G., Di Salvo, J., Kilburn, R., Cascieri, M.A., Kurtz, M.M., Tsao, K.-L., Gould, S.L., Chicchi, G.G., 2003. Comparison of the functional blockade of rat substance P (NK_1) receptors by GR205171, RP67580, SR140333 and NKP-608. Neuropharmacology 45, 231–241.

Schloss, P., Henn, F., 2004. New insight into the mechanisms of antidepressant therapy. Pharmacol. Ther. 102, 47–60.

Segal, R.A., Greenberg, M.E., 1996. Intracellular signaling pathways activated by neurotrophic factors. Annu. Rev. Neurosci. 19, 463–489.

Spigset, O., Martensson, B., 1999. Fortnightly review: drug treatment of depression. BMJ 318, 1188–1191.

Stout, S.C., Owens, M.J., Nemeroff, C.B., 2001. Neurokinin 1 receptor antagonists as potential antidepressants. Annu. Rev. Pharmacol. Toxicol. 41, 877–906.

Sweatt, J.D., 2004. Mitogen-activated protein kinases in synaptic plasticity and memory. Curr. Opin. Neurobiol. 14, 311–317.

Thomas, G.M., Huganir, R.L., 2004. MAPK cascade signalling and synaptic plasticity. Nat. Rev., Neurosci. 5, 173–183.

Tohda, M., Qi, Z., Watanabe, H., 1999. Influence of chronic treatment with imipramine on mRNA levels in rat brain: elevation of glyceraldehyde-3-phosphate dehydrogenase levels. Jpn. J. Pharmacol. 81, 393–396.

Tiraboshi, E., Tardito, D., Kasahara, J., Moraschi, S., Pruneri, P., Gennarelli, M., Racagni, G., Popoli, M., 2004. Selective phosphorylation of nuclear CREB by fluoxetine is linked to activation of CaM kinase IV and MAP kinase cascades. Neuropsychopharmacology 29, 1831–1840.

Valjent, E., Pagès, C., Hervé, D., Girault, J.-A., Caboche, J., 2004. Addictive and non-addictive drugs induce distinct and specific patterns of ERK activation in mouse brain. Eur. J. Neurosci. 19, 1826–1836.

van der Hart, M.G., Czeh, B., de Biurrun, G., Michaelis, T., Watanabe, T., Natt, O., Frahm, J., Fuchs, E., 2002. Substance P receptor antagonist and clomipramine prevent stress-induced alterations in cerebral metabolites, cytogenesis in the dentate gyrus and hippocampal volume. Mol. Psychiatry 7, 933–941.

Weitzdoerfer, R., Fountoulakis, M., Lubec, G., 2001. Aberrant expression of dihydropyrimidinase related proteins-2, -3 and -4 in fetal Down syndrome brain. J. Neural Transm., Suppl. 61, 95–107.

Whitnall, M.H., 1993. Regulation of the hypothalamic corticotropin-releasing hormone neurosecretory system. Prog. Neurobiol. 40, 573–629.

Xu, H., Richardson, J.S., Li, X.-M., 2003. Dose-related effects of chronic antidepressants on neuroprotective proteins BDNF, Bcl-2

and Cu/Zn-SOD in rat hippocampus. Neuropsychopharmacology 28, 53–62.

Yoshimura, Y., Shinkawa, T., Taoka, M., Kobayashi, K., Isobe, T., Yamamuchi, T., 2002. Identification of protein substrates of Ca2+/calmodulin-dependent protein kinase II in the post-synaptic density by protein sequencing and mass spectrometry. Biochem. Biophys. Res. Commun. 290, 948–954.

Zobel, A.W., Nickel, T., Künzel, H.E., Ackl, N., Sonntag, A., Ising, M., Holsboer, F., 2000. Effects of the high-affinity corticotropin-releasing hormone receptor 1 antagonist R121919 in major depression: the first 20 patients treated. J. Psychiatr. Res. 34, 171–181.

Anal Bioanal Chem (2008) 390:1327–1342
DOI 10.1007/s00216-008-1837-y

PAPER IN FOREFRONT

Application of partial least squares discriminant analysis and variable selection procedures: a 2D-PAGE proteomic study

Emilio Marengo · Elisa Robotti · Marco Bobba · Alberto Milli · Natascia Campostrini · Sabina Carla Righetti · Daniela Cecconi · Pier Giorgio Righetti

Received: 12 November 2007 / Revised: 21 December 2007 / Accepted: 8 January 2008 / Published online: 29 January 2008

Abstract 2D gel electrophoresis is a tool for measuring protein regulation, involving image analysis by dedicated software (PDQuest, Melanie, etc.). Here, partial least squares discriminant analysis was applied to improve the results obtained by classic image analysis and to identify the significant spots responsible for the differences between two datasets. A human colon cancer HCT116 cell line was analyzed, treated and not treated with a new histone deacetylase inhibitor, RC307. The proteins regulated by RC307 were detected by analyzing the total lysates and nuclear proteome profiles. Some of the regulated spots were identified by tandem mass spectrometry. The preliminary data are encouraging and the protein modulation reported is consistent with the antitumoral effect of RC307 on the HCT116 cell line. Partial least squares discriminant analysis coupled with backward elimination variable selection allowed the identification of a larger number of spots than classic PDQuest analysis. Moreover, it allows the achievement of the best performances of the model in terms of prediction and provides therefore more robust and reliable results. From this point of view, the multivariate procedure applied can be considered a good alternative to standard differential analysis, also taking into account the interdependencies existing among the variables.

Keywords Two-dimensional maps · Proteomics · Multivariate statistical methods · Partial least squares discriminant analysis · Variable selection procedures

Abbreviations

BE	backward elimination
CHAPS	3-[(3-cholamidopropyl)dimethylammonium-1-propanesulfonate
DA	discriminant analysis
HDAC	histone deacetylase
HSPA5	heat shock 70-kDa protein 5
IEF	isoelectric focusing
IPG	immobilized pH gradient
LDA	linear discriminant anlysis
LV	latent variable
MALDI-TOF-TOF	matrix-assisted laser desorption/ionization tandem time-of-flight
PAGE	polyacrylamide gel electrophoresis

E. Marengo (✉) · E. Robotti · M. Bobba
Dipartimento di Scienze dell'Ambiente e della Vita,
Università degli Studi del Piemonte Orientale,
Via Bellini 25/G,
15100 Alessandria, Italy
e-mail: marengoe@tin.it

A. Milli · D. Cecconi
Dipartimento Scientifico e Tecnologico,
Laboratorio di Proteomica, Università degli Studi di Verona,
Strada le Grazie 15,
37134 Verona, Italy

N. Campostrini
Dipartimento di Medicina Clinica e Sperimentale,
Sezione di Medicina Interna B, Università degli Studi di Verona,
P.le L.A. Scuro 10,
37134 Verona, Italy

S. C. Righetti
Istituto Nazionale per lo Studio e la Cura dei Tumori,
Via Venezian 1,
20133 Milano, Italy

P. G. Righetti
Dipartimento di Chimica,
Materiali e Ingegneria Chimica "Giulio Natta",
Politecnico di Milano, Via Mancinelli 7,
20131 Milano, Italy

PCA	principal component analysis
PLS	partial least squares
PRDX1	peroxiredoxin I
RMSE	root mean square error
RMSEC	root mean square error in calibration
RMSECV	root mean square error in cross-validation
RMSEP	root mean square error in prediction
RP-HPLC-ESI MS/MS	reversed-phase high-performance liquid chromatography electrospray ionization tandem mass spectrometry
SDS	sodium dodecyl sulfate
SIMCA	soft-independent model of class analogy
SSP	standard spot
STMN1	stathmin 1
Tris	tris(hydroxymethyl)aminomethane

Introduction

It is common knowledge that 2D gel electrophoresis is one of the most widespread techniques for outlining cellular protein profiles. It achieves a double separation of proteins according to their isoelectric point (first dimension) and molecular mass (second dimension), thus providing a final 2D map [2D polyacrylamide gel electrophoresis (PAGE)] where the proteins are represented as spots spread over the gel matrix. Each spot on the final map represents a protein or a group of similar proteins that cannot be separated further.

The large amount of information contained in groups of maps, belonging to the same specimen or to specimens to be compared (e.g., control and pathological subjects), is usually extracted by means of dedicated software packages (PDQuest, Melanie, Progenesis, etc.). The classic differential analysis thus carried out is based on the direct analysis of images of maps, previously scanned via a densitometer. The multistep procedure consists briefly of (1) background correction of the maps to be analyzed, (2) alignment of the maps, (3) identification of the spots present independently on each map (fitting with Gaussian distributions), (4) matching of the spots present on all the maps belonging to the same group (e.g. control or treated samples) and (5) achievement of the "master gel" for each class containing the spots present contemporarily on all the maps belonging to the same group.

The final differential analysis can be carried out using the same software, which will ultimately list the spots responsible for the differences between the master gels of each group of samples. As an alternative, a spot volume dataset can be obtained where each sample (map) of each group is described in terms of the volumes (sum of optical densities) of all the spots identified, namely, those contained on all the maps of that group contemporarily and those contained in a restricted subset of maps. Such datasets, where a certain number of samples are described by a large number of variables (spots), can be effectively treated by means of multivariate statistical tools, such as principal component analysis (PCA), classification methods [linear discriminant analysis (LDA), soft-independent model of class analogy, (SIMCA)] and cluster analysis techniques. Multivariate tools are widely applied to proteomic datasets [1–9]. Some recent applications of PCA, SIMCA and other multivariate tools were presented also by our research group [10–13]. Other papers have reported the use of partial least squares (PLS) DA as a classification tool for identifying the differences between groups of samples [14–21] or the use of chemometric tools for protein selection in gel electrophoresis [22]. In the present report, PLS-DA was applied coupled with variable selection procedures (backward elimination, BE) for identifying the actually significant spots responsible for the differences existing between two groups of samples in two different datasets: nuclei from HCT116 (six control and five treated samples) and total cell lysates from HCT116 (five control and five treated samples). The results obtained from the statistical analyses were then compared with those provided by a classic PDQuest differential analysis based on individual Student *t* tests (α=0.05). The treatment of the HCT116 cell line was performed by using a new inhibitor of histone deacetylase (HDAC), named RC307. HDACs are zinc-dependent metalloenzymes that catalyze the hydrolysis of acetylated lysine residues. Most HDAC inhibitors have a zinc binding group, such as hydroxamic acid. During the last decade, a number of HDAC inhibitors have been identified that induce cultured tumor cells to undergo growth arrest, differentiation and/or apoptotic cell death [23]. For example, we have reported, using a proteomic profiling with 2D maps, that the hydroxamic acid derivative HDAC inhibitor trichostatin A induces apoptosis and G2-phase cell cycle arrest in pancreatic cancer cell lines [24, 25]. Recently, we have found that a novel series of HDAC inhibitors, still belonging to the hydroxamic acid family, exhibit a broad-spectrum inhibition profile, characterized by a marked effect on acetylation of non-histone proteins. The novel inhibitor tested here (RC307, patent applied for) belongs to this last class.

The PLS-DA analysis applied in this study enabled us to validate and improve the PDQuest analysis, while obtaining also information about the mechanism of action of this new antitumoral agent. In fact, some of the significant differentially expressed spots (detected by PDQuest coupled with PLS-DA or by the PLS-DA procedure with variable selection)

were further identified by matrix-assisted laser desorption/ionization tandem time-of-flight (MALDI-TOF-TOF) analysis and reversed-phase high-performance liquid chromatography electrospray ionization tandem mass spectrometry (RP-HPLC-ESI MS/MS) and some hypotheses were made on the biological role played by the proteins identified.

Theory

Partial least squares discriminant analysis

PLS [26–30] is a multivariate regression method establishing a relationship between one or more dependent variables (Y) and a group of descriptors (X). X and Y variables are modeled simultaneously, to find the latent variables (LVs) in X that will predict the LVs in Y. These LVs (also called PLS components) are similar to the principal components calculated from PCA [26, 27]: they are computed hierarchically, i.e., the first component accounts for the largest amount of information, followed by the other components, accounting for the maximum residual variance. If several responses are presents (more than one Y variable), they are modeled together by the so-called PLS2 algorithm. Since LVs are computed hierarchically, the last LVs are mostly responsible for random variations and experimental error. The optimal number of LVs, i.e., those modeling information in X useful to predict the response Y but avoiding overfitting, is determined on the basis of the residual variance in prediction. Here, leave-one-out cross-validation is applied to evaluate the predictive ability and select the optimal number of LVs on X and Y contained in the final model. Owing to the small number of samples and the large number of variables, the final model can be validated only by leave-one-out cross-validation. However, leave N out (N ranging from 2 to 3) was also applied to verify the existence of a significant model: this procedure provided results similar to leave-one-out cross-validation.

In the case where a large number of descriptors (X variables) are present or a large experimental error is expected, it can be quite difficult to obtain a final model with a suitable predictive ability. In these cases, techniques for variable selection are usually exploited. Here, two subsequent strategies were applied: a first simplification of the model by eliminating groups of non significant X variables, reducing by about 50% the number of variables present, according to the minimum error in cross-validation; a second phase where variables were eliminated one at a time to provide the final model with the overall minimum error in cross-validation (BE procedure).

PLS was originally set up to model continuous responses but it can be applied even for classification purposes by establishing an appropriate Y related to the association of each sample to a class. Since two classes are present in this case (control and treated samples, in two different datasets), a binary Y variable is added to each dataset, coded so that −1 is attributed to control samples and +1 to treated samples. The regression is then carried out between X-block variables (spot volumes) and the Y just established. This application for classification purposes is called PLS-DA.

Evaluation of models

The coefficient of multiple determination, R^2, for PLS was calculated as

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}, \quad (1)$$

where the two sums run on the samples used for calibration (R^2), or for validation (R^2_{CV}), $\hat{y}_i$ is the predicted value of the response for the ith experiment and $\bar{y}$ is the average response of the samples used for calibration (R^2), or for validation (R^2_{CV}).

The root mean square error (RMSE) is calculated as

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n}}, \quad (2)$$

where the sum runs on the samples used for calibration (root mean square error in calibration, RMSEC), for validation (root mean square error in cross-validation, RMSECV) or for genuine prediction (root mean square error in prediction, RMSEP), and $\hat{y}_i$ is the predicted value of the response for the ith experiment.

In this case, the best model complexity in regard to both the number of X variables present and the number of LVs in the model is selected by the minimum value of the RMSECV (leave-one-out procedure). The RMSEP was not calculated since no separation of the samples in training, evaluation and production sets was possible, owing to the small number of samples present in both datasets.

Materials and methods

Cell culture and drug exposure

The human colon cancer HCT116 cell line was grown and maintained as monolayers in RPMI-1640 medium supplemented with 10% fetal calf serum (Invitrogen, Carlsbad, CA, USA). Exponentially growing cells were seeded in

150-cm^2 flasks and, 24 h later, they were exposed to the new HDAC inhibitor for 24 h (IC_{80}=2.5 μM). After drug exposure, cells (10^9 for nuclear extraction and 10^8 for total lysates) were harvested, pelleted and stored at −80 °C.

2D gel electrophoresis

Total protein extraction from 10^8 cells not treated and treated with the HDAC inhibitor RC307 was performed in a 2D solubilizing/lysing solution: 7 M urea (Sigma-Aldrich, St. Louis, MO, USA), 2 M thiourea (Sigma), 3% 3-[(3-cholamidopropyl)dimethylammonium-1-propanesulfonate (CHAPS; Sigma), 20 mM tris(hydroxymethyl)aminomethane (Tris; Sigma), 1% pH 3–10 Ampholine (Fluka, Buchs, Switzerland) and one protease inhibitor cocktail tablet (Complete, Mini; Roche, Basel, Switzerland). The samples were then sonicated five times for 30 s on ice with 1-min rest in-between and the sonicates were centrifuged for 10 min at 10,000*g* at 4 °C to remove the nucleic acids complexed with ampholytes. Additionally, a nuclear protein extraction from 10^9 cells not treated and treated with the HDAC inhibitor RC307 was obtained with the CelLytic NuCLEAR extraction kit (Sigma) by following the manufacturer's instructions. Then 1% pH 3–10 Ampholine was added and the samples were centrifugated as described above, for removal of nucleic acids. Both total cell lysates and nuclear extracts were incubated with 5 mM tributyl phosphine and 20 mM acrylamide for 60 min at room temperature to reduce protein disulfide bonds and alkylate the cysteine thiolic groups. The reaction was blocked by the addition of 10 mM dithiothreitol (Sigma) and the samples were collected and stored at −80 °C. Protein concentration was evaluated with DC protein assay (Bio-Rad, Hercules, CA, USA) based on the Lowry method. Seventeen-centimeter-long, pH 3–10 immobilized pH gradient (IPG) strips (Bio-Rad) were rehydrated for 8 h with 450 μL of 2D solubilizing solution (7 M urea, 2 M thiourea, 3% CHAPS and 20 mM Tris) containing 3 mg/mL of total protein from cells. Isoelectric focusing (IEF) was carried out with a Protean IEF cell (Bio-Rad), with a initial linear voltage gradient from 0 to 1,000 V in 15 h (overnight) and then by applying an exponential voltage gradient up to 10,000 V for about 5 h, with a limiting current of 50 μA per strip. The applied product of time and voltage was 70,000 Vh for each strip and the temperature was set at 20 °C. For the second dimension, the IPG strips were equilibrated for 26 min by rocking them in a solution of 6 M urea, 2% sodium dodecyl sulfate (SDS), 20% glycerol, 375 mM Tris-HCl, pH 8.8. The IPG strips were then laid on an 8–18% T gradient SDS-PAGE gel with 0.8% agarose in Tris/glycine/SDS running buffer (192 mM glycine, 0.1% SDS and Tris to pH 8.3). The second dimension was performed in a Protean Plus Dodeca cell (Bio-Rad) with Tris/glycine/SDS running buffer. The electrophoresis was conducted with continuous cooling and mixing by setting a current of 40 mA for each gel for 3 min, then 2 mA per gel for 1 h, and 20 mA per gel until the track dye, bromophenol blue, reached the anodic end of the gels. The protein zones were finally revealed with Sypro Ruby stain (Bio-Rad). Gels were incubated in a fixing solution containing 40% ethanol and 10% acetic acid for 30 min followed by overnight staining in a ready-to-use Sypro Ruby solution. Destaining was performed in 10% methanol and 7% acetic acid for 1 h, followed by a rinse of at least 3 h in pure water.

Protein pattern differential analysis

2D gel analysis was performed by PDQuest software (Bio-Rad), version 7.3. Each gel was analyzed for spot detection, background subtraction and protein spot optical density intensity quantification (spot quantity definition). The gel image showing the highest number of spots and the best protein pattern was chosen as a reference template, and spots in a standard gel were then matched across all gels. Spot quantity values were normalized in each gel by dividing the raw quantity of each spot by the total quantity of all the spots included in the standard gel. Two distinct differential analyses were performed, one for total lysates and one for nuclear extracts. In both analyses the gels were divided into two separated groups (control and RC307-treated samples) and, for each protein spot, the average spot quantity value and its variance coefficient in each group were determined. A Student *t* test was performed in order to compare the two groups and identify sets of proteins that showed a statistically significant difference with a confidence level of 0.05.

In-gel digestion

Spots were cut out from 2D Sypro Ruby stained gels and subjected to in-gel trypsin digestion. Briefly, spots were destained [one 15-min 300-μL wash in 100 mM NH_4HCO_3; one 15-min 300-μL wash in 50% 100 mM NH_4HCO_3 (v/v), 50% acetonitrile; one 5-min wash in 100% acetonitrile] and dried at 37 °C. The gel pieces were swollen in 10 μL of a digestion buffer containing 100 mM NH_4HCO_3 and 40 ng/μL trypsin (modified porcine trypsin, sequencing grade, Promega, Madison, WI, USA). After 10 min, 40 μL of 100 mM NH_4HCO_3 was added to the gel pieces and digestion was allowed to proceed at 37 °C overnight. The supernatants were collected and the peptides were extracted in an ultrasonic bath for 10 min (twice with 50 μL of 50% acetonitrile, 50% H_2O with 1% formic acid v/v; once with 25 μL of acetonitrile), whereby the supernatant was added every time to the previous one. Tryptic peptides were dried by vacuum centrifugation and redissolved in 20 μL of 0.1% formic acid in water and

purified by using Zip-Tip C18 (Millipore Bedford, MA, USA).

Peptide sequencing by MALDI-TOF-TOF and RP-HPLC-ESI MS/MS

The lyophilized tryptic peptides were dissolved in 10 mg/mL α-cyano-4-hydroxycinnamic acid containing 0.1% trifluoroacetic acid and 50% acetonitrile. Mass spectrometry analysis was conducted with a MALDI-TOF-TOF mass spectrometer 4800 proteomics analyzer (Applied Biosystems, Framingham, MA, USA). Data were analyzed using GPS Explorer software (Applied Biosystems) and MASCOT software (Matrix Science, London, UK). NCBInr and human were selected as the database and taxonomy, respectively. The peptide mixtures were also analyzed by using a nanoflow high-performance liquid chromatography system (Ultimate; Switchos; Famos; LC Packings, Amsterdam, The Netherlands) coupled to an Esquire 300 plus ion trap (Bruker-Daltonik, Germany). A sample volume of 10 μL was loaded by the autosampler onto a homemade 2-cm fused-silica precolumn (75-μm inner diameter, 375-μm outer diameter; Resprosil C18-AQ, 3 μm, Ammerbuch-Entringen, Germany) at a flow rate of 2 μL/min. Sequential elution of peptides was accomplished using a flow rate of 200 nL/min and a linear gradient from solution A (2% acetonitrile, 0.1% formic acid) to 50% of solution B (98% acetonitrile, 0.1% formic acid) in 40 min over the precolumn in-line with a homemade 10–15-cm resolving column (75-μm inner diameter, 375-μm outer diameter; Resprosil C18-AQ, 3 μm, Ammerbuch-Entringen, Germany). Peptides were eluted directly into an Esquire 3000 plus ion trap (Bruker-Daltonik, Germany). Data were analyzed using MASCOT software (Matrix Science, London, UK). NCBInr and human were selected as the database and taxonomy, respectively.

Datasets

The analysis was carried out independently for the two cases under investigation, i.e., nuclei and lysates datasets. The nuclei dataset consisted of 11 samples (six control and five treated samples) described by 779 spot volumes. The dataset deriving from lysates consisted of ten samples (five control and five treated samples) described by 528 spots. For both datasets, a binary Y variable was added, representing the belonging of each sample to a class: −1 was attributed to control samples and +1 to treated samples. For both cases, each spot is univocally identified by its standard spot (SSP) number.

Software

Differential analysis was performed using PDQuest (version 7.3, Bio-Rad). Multivariate analysis and graphical representations were performed using Statistica version 7.1 (Statsoft, USA), The Unscrambler version 9.5 (CAMO, Norway), Parvus 2006 (kindly provided by Michele Forina, University of Genova, Italy) and Microsoft Excel (Microsoft, USA).

Results and discussion

The human colon cancer HCT116 cell line treated with the new HDAC inhibitor RC307 was chosen as a model for comparing the classic PDQuest image analysis with PLS-DA coupled with variable selection procedures. The study was performed by measuring differential protein expression in the total lysates and nuclei of control and treated cells. In particular, by PDQuest analyses, we detected a total of 48 spots regulated in lysates (27 upregulated and 21 downregulated) and a total of 46 spots regulated in nuclei (24 upregulated and 22 downregulated). Representative standard 2D maps of the two datasets (lysates and nuclei) are reported in Fig. 1a and b, respectively.

The spot volume datasets provided by PDQuest were then analyzed by BE-PLS-DA to identify the upregulated and downregulated spots in both datasets. The results obtained were compared with those from classic PDQuest differential analysis.

As already pointed out in “Theory,” variable selection strategies can be applied to regression and classification problems to identify the best subset of variables providing a model characterized by the best predictive performance.

In this case, the BE algorithm was applied to both datasets, following a two steps procedure:

1. Groups of variables were first eliminated; in each step no more than 6% of the predictors were eliminated. A first reduction of the dataset of more than 50% was finally obtained.
2. Predictors were then eliminated one at a time. The model containing the number of predictors able to provide the minimum error in cross-validation (RMSECV) was selected.

Nuclei dataset

Backward elimination partial least squares discriminant analysis

The dataset of dimensions 11×779 (11 being the samples and 779 being the variables: 778 spot volumes and the Y variable) was autoscaled before PLS-DA. The autoscaling procedure is particularly important in proteomic datasets where large scale-effects are present among the variables, allowing all variables to account for the same amount of variance. In this way also differences between groups due to variations in small spots can be effectively detected.

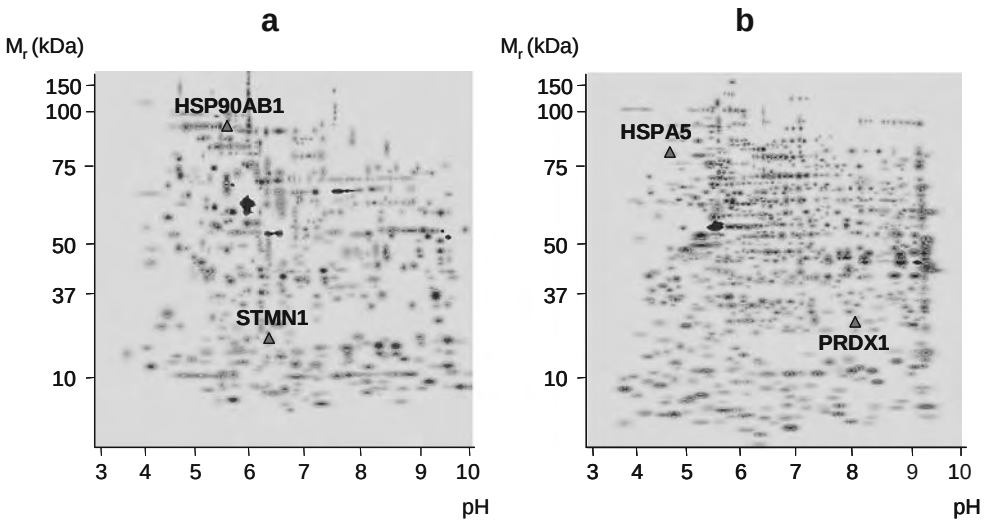

Fig. 1 Examples of standard 2D maps of total cell lysates (**a**) and nuclei (**b**) of the HCT116 cell line. The proteins identified are marked by a *triangle* and the corresponding gene name. The pH range of 3–10 in the first dimension was linear. The acrylamide gradient in the second dimension was 8–18% T

The first screening by the BE procedure allowed us to reduce the dataset to 207 predictors. Then predictors were eliminated one at a time: the best predictive ability was obtained with a final model containing 61 variables and the first LV.

The percentage of explained and cumulative explained variance of the *Y* and *X* variables are reported in Table 1 for the first nine LVs calculated. The first LV was retained as significant since it explains about 80% of the variance contained in the *X* block and more than 99% on *Y*.

The score and loading plots of the first two LVs are represented in Fig. 2a. The score plot represents the samples well separated in control samples (at negative values on LV_1, represented as circles) and treated samples (at positive values on the same LV, represented as squares). The analysis of the loading plot allowed us to identify at positive loadings on LV_1 the spots more intense in treated samples (upregulated after the treatment) and at negative loadings on LV_1 the spots more intense in control samples (downregulated after the treatment).

The regression coefficients calculated on the basis of the first significant LV are represented in Fig. 3a on a "virtual" map on a color scale. Each spot is represented as a circle centered on the *x*–*y* position (identified by PDQuest) and identified by its SSP number. Spots are represented on a color scale according to their increasing positive (from light to dark red) or negative (from light to dark blue) coefficient. Spots characterized by large negative coefficients are thus more intense in control samples (downregulated after the treatment), while spots characterized by large positive coefficients are more intense in treated samples (upregulated after the treatment).

Table 2 reports the performances of the model built with 61 predictors and the first LV as the significant one, both in fitting and in prediction.

The model shows very good performances regarding both its ability to fit data and its capability in terms of prediction,

Table 1 Percentage variance explained by the latent variables (*LV*) calculated for *X* and *Y* variables for the nuclei dataset and the lysates dataset

	Nuclei dataset					Lysates dataset			
	BE-PLS-DA		PDQuest PLS-DA			BE-PLS-DA		PDQuest PLS-DA	
	X	*Y*	*X*	*Y*		*X*	*Y*	*X*	*Y*
LV 1	80.18	99.99	68.12	99.46	LV 1	50.93	99.88	60.76	97.56
LV 2	5.89	0.01	8.71	0.38	LV 2	7.13	0.10	6.24	2.11
LV 3	3.04	0.002	7.58	0.12	LV 3	6.21	0.01	5.30	0.26
LV 4	1.57	0.0001	3.60	0.034	LV 4	6.83	0.001	4.50	0.06
LV 5	2.24	0.00002	1.81	0.011	LV 5	5.26	0	4.90	0.01
LV 6	1.33	0	2.63	0.002	LV 6	5.60	0	5.68	0.001
LV 7	1.16	0	2.27	0	LV 7	5.49	0	4.55	0
LV 8	1.67	0	1.13	0	LV 8	6.38	0	4.76	0
LV 9	1.60	0	2.14	0					

BE backward elimination, *PLS-DA* partial least squares discriminant anaylsis

according to the values of R^2 calculated. Regarding the error produced by the model (RMSE), it is extremely good in fitting (about 0.0107) and very good in prediction (about 0.0192). These results are also confirmed by the representation of the calculated versus the measured responses for both calibration and validation samples (Fig. 4a), identifying extremely low variations along the *y*-axis.

PLS-DA on PDQuest spots: comparison with the BE-PLS-DA model

For comparison, a PLS-DA model was built considering the 46 spots identified as the relevant ones by PDQuest as the *X* block and the *Y* variable already coded for the previous case as the dependent variable. Data were autoscaled before performing PLS-DA. Table 1 reports the percentage of explained variance for the first nine LVs calculated, compared with that explained by the first nine LVs calculated by BE-PLS-DA. The first LV appears to be the significant one since it explains more than 99% of the information contained in the *Y* variable but only 68% of that in the *X* block. Figure 2b shows score and loading plots for the first two LVs. As for the previous case, the two groups of samples are clearly separated along the first LV: control samples (circles) at negative values and treated samples (squares) at positive scores along LV_1. The corresponding loading plot identifies 22 spots at negative values along LV_1, more intense in control samples (downregulated after the treatment), and 24 spots at positive values along the same LV, more intense in treated samples (upregulated after the treatment).

The coefficients calculated with the first LV in the final model are represented in Fig. 3b on a "virtual" map on a

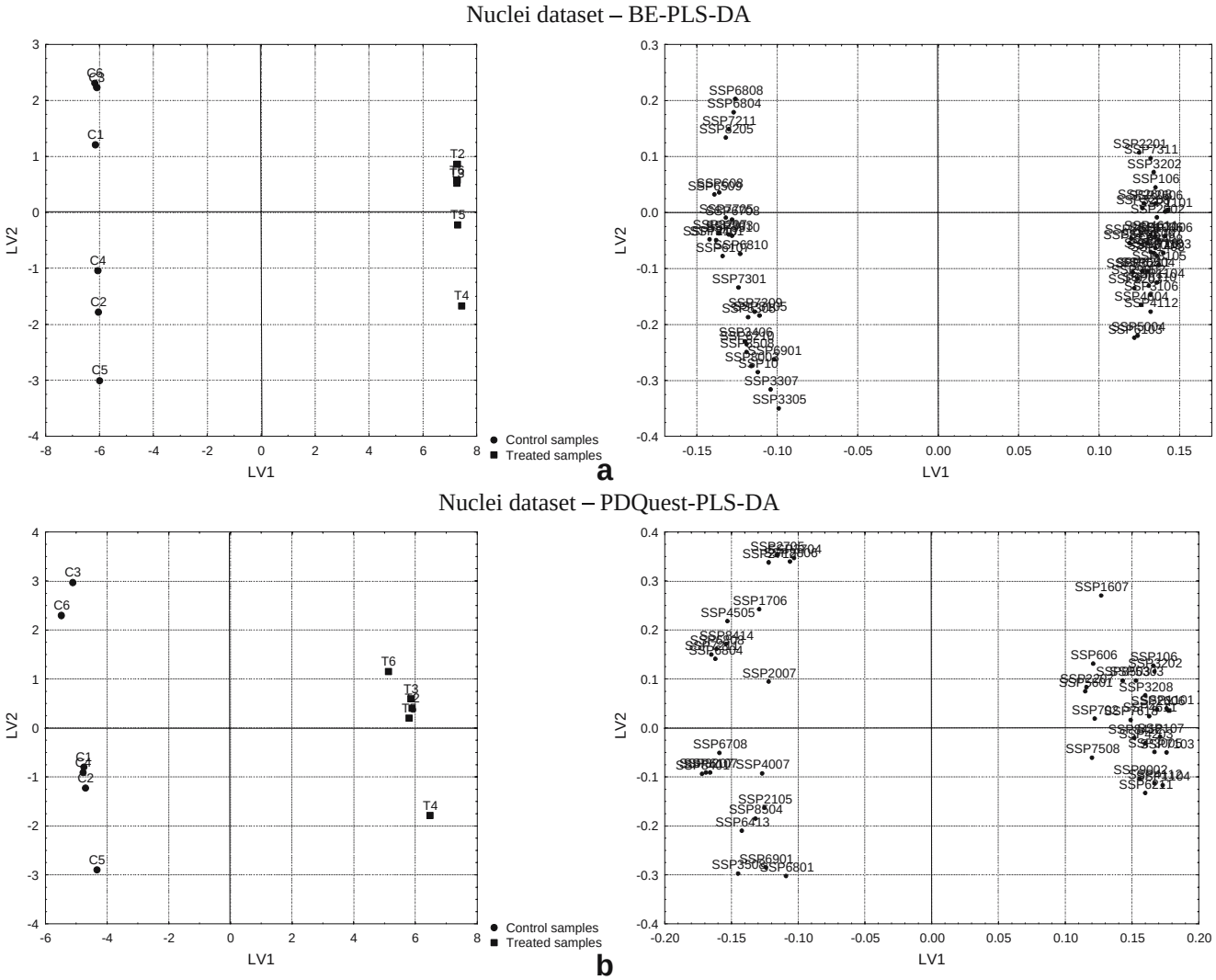

Fig. 2 Nuclei dataset: score plot and loading plot (**a**) of the first two latent variables for the final backward elimination partial least squares discriminant analysis (*BE-PLS-DA*) model with variable selection; score plot and loading plot (**b**) of the first two latent variables for the PLS-DA model containing the spots identified by PDQuest. For the score plots control samples are represented as *circles* and treated samples as *squares*; for the loading plots *X* variables are represented as *circles* and the *Y* variables as *squares*. *LV* latent variable

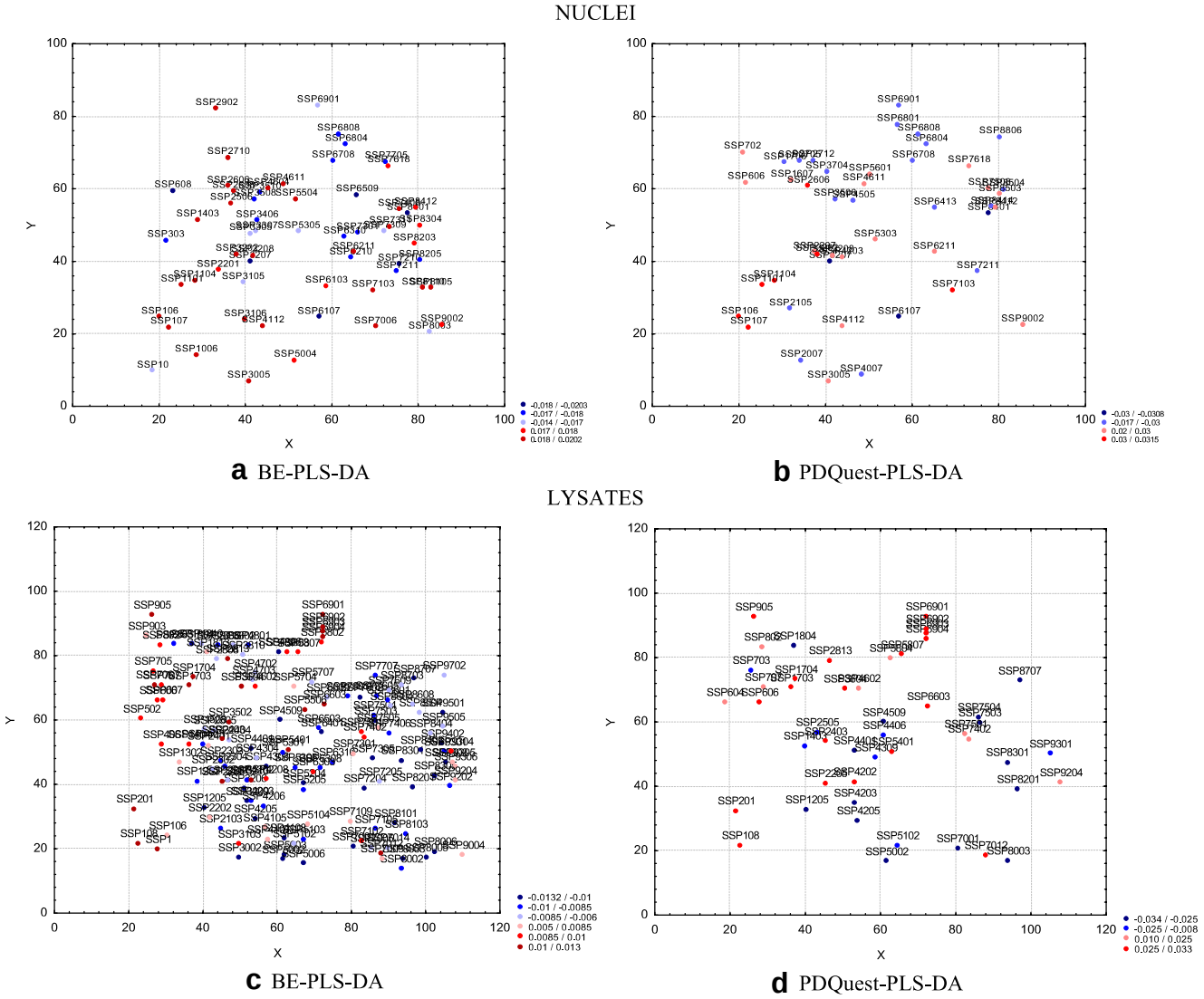

Fig. 3 Nuclei dataset: regression coefficients represented for the final BE-PLS-DA (**a**) and PDQuest PLS-DA (**b**) models. Lysates dataset: regression coefficients represented for the final BE-PLS-DA (**c**) and PDQuest PLS-DA (**d**) models. Each spot is represented on a virtual map as a *circle* centered on the *x*–*y* position identified by PDQuest. The coefficients are represented on a color scale: from light to dark blue corresponds to increasing negative coefficients; from light to dark red corresponds to increasing positive coefficients

Table 2 Performance of BE-PLS-DA and PDQuest PLS-DA models in fitting and prediction for nuclei and lysates datasets (one LV in all models)

	BE-PLS-DA		PLS-DA on PDQuest	
	Calibration	Validation	Calibration	Validation
Nuclei dataset				
R^2	0.9999	0.9999	0.9946	0.9888
RMSE	0.01070	0.01925	0.08114	0.1091
Lysates dataset				
R^2	0.9988	0.9980	0.9756	0.9537
RMSE	0.03846	0.04549	0.1745	0.2253

RMSE root mean square error

color scale, as for the previous case. From the results reported in Table 2, the performances of this model appear good, according to the values of R^2 calculated both in fitting and in validation; the same conclusions can be drawn even if the error produced by the model is examined. However, if the performances of this last model are compared with those of the BE-PLS-DA model, it is clear that some useful information is not taken into account by PDQuest differential analysis. The lower values of R^2 in fitting (and the consequent larger value of RMSE for calibration) show that the PLS-DA model on PDQuest spots is characterized by a smaller fitting ability. The performances of this last model appear even worse if the results for validation are considered, and we obtained a final error about 5 times larger than that calculated by the

BE-PLS-DA model. The same conclusions can be drawn from Fig. 4b, representing the calculated versus the measured responses both in fitting and in prediction: larger deviations can be identified along the *y*-axis if comparisons are made with the BE-PLS-DA model (Fig. 4a).

The comparison between the significant spots identified by PDQuest differential analysis and those identified by BE-PLS-DA shows that the PLS-DA procedure with variable selection identifies about 52.17% of the spots identified also by PDQuest (24 of 46); some of them are identified exclusively by PDQuest (47.83%: 22 of 46) and 37 spots are identified only by means of BE-PLS-DA. BE-PLS-DA allows us to obtain a larger number of significant spots with respect to standard differential analysis. At first sight, the two procedures (PLS-DA with variable selection and PDQuest differential analysis) seem complementary since BE-PLS-DA does not identify as relevant some spots identified by PDQuest and vice versa; it must be stressed, however, that the choice of the best set of significant spots made by BE-PLS-DA ensures the best performances in prediction, while this is not the case with differential analysis based on *t* tests. From this point of view the results provided by BE-PLS-DA appear more robust and reliable; furthermore, BE-PLS-DA provides a set of spots taking into consideration the correlation structure of the dataset, thus considering the interdependencies existing between the spots identified. The spots identified by the two procedures, divided into upregulated and downregulated after the treatment, together with the values of the coefficients calculated by the BE-PLS-DA model, are reported in Table 3, according to their decreasing positive or negative coefficient. Table 3 also reports the PLS-DA coefficients

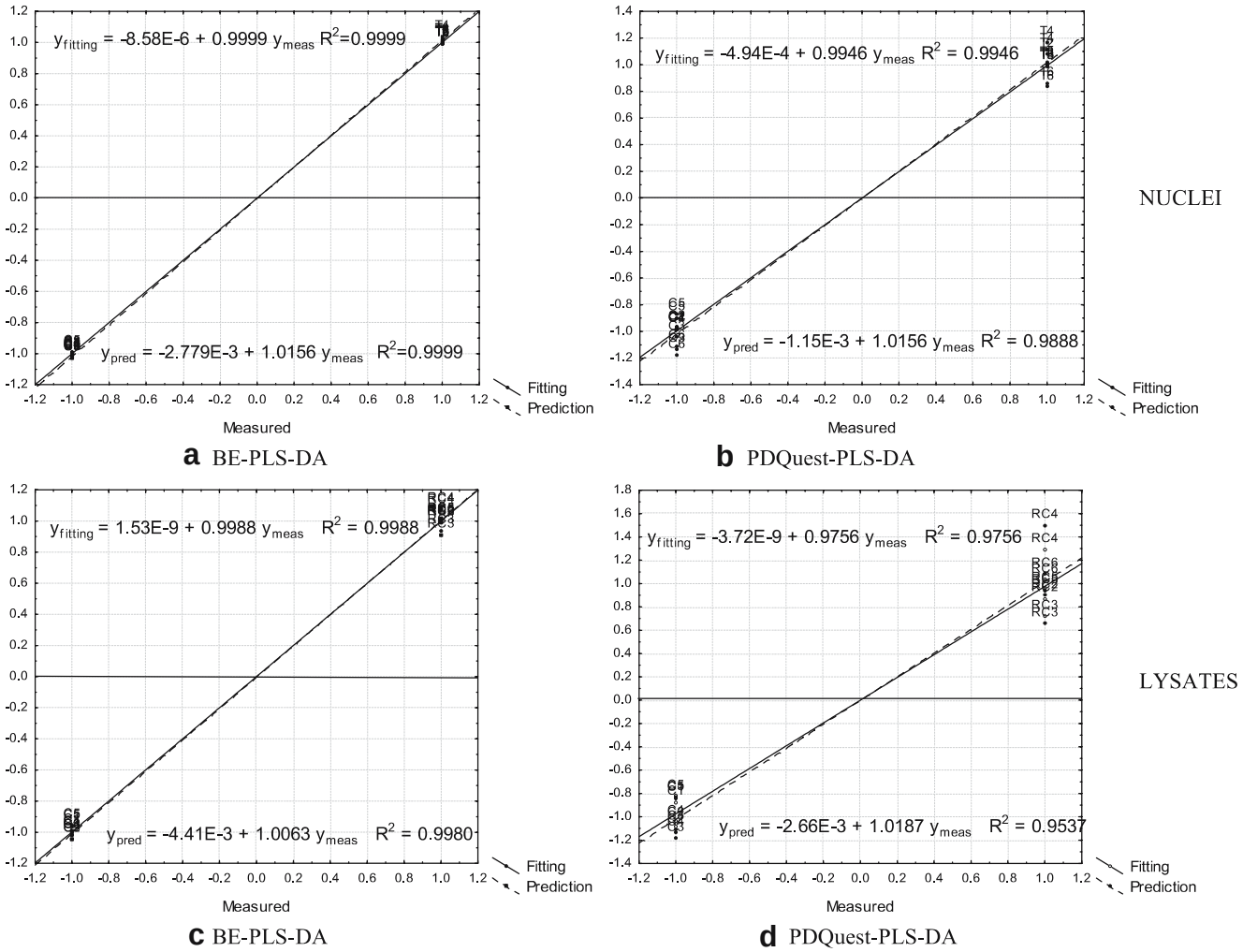

Fig. 4 Calculated *Y* versus measured *Y* in fitting and prediction (leave-one-out cross-validation): for the nuclei dataset for the final PLS-DA model with variable selection (one latent variable in the model) (**a**) and for the PLS-DA model based on the spots identified by PDQuest (**b**); for the lysates dataset for the final PLS-DA model with variable selection (one latent variable in the model) (**c**) and for the PLS-DA model based on the spots identified by PDQuest (**d**). The calibration samples are represented as *circles*, the validation samples as *squares*

for the model built on the spots identified as significant by PDQuest, together with the significance levels for each variable in each model. The significance levels were calculated by a *t* test on each coefficient in cross-validation: values less than 0.05 indicate variables that have a probability level of more than 95% and are considered significant. The results show that the coefficients calculated by the BE-PLS-DA model are all significant and their values are similar to those provided by PLS-DA for the spots identified by PDQuest. However, the BE-PLS-DA procedure includes more variables that show a high relevance.

Lysates dataset

Backward elimination partial least squares discriminant analysis

The dataset deriving from lysate analysis, of dimensions 10×529 (10 being the samples and 529 the variables: 528 spot volumes and one *Y* variable), was autoscaled before the application of PLS-DA. The first step of the BE procedure provided a dataset of 227 predictors. Then, variables were eliminated one at a time to identify the model with the best predictive ability. The final model contained 146 predictors and only the first LV. Table 1 reports the percentage of explained and cumulative explained variance for both *Y* and *X* variables for the first eight LVs calculated.

The first LV was considered the relevant one, since it accounts for more than 99% of the variance contained on the *Y* variable, even if it is related to only 50% of variance on the *X* block.

Figure 5a shows the score plot and loading plot of the first two LVs calculated. From the score plot, it is clear that the first LV is effective in the separation of the two groups of samples: control samples (circles) at negative values and treated samples (squares) at positive ones. The corresponding loading plot allows the identification of two groups of spots: the spots present at negative values along LV_1 are more intense in control samples (downregulated after the treatment), while those at positive loadings on the same LV are more intense in treated samples (upregulated after the treatment).

Figure 3c shows the regression coefficients calculated on the basis of the first LV in the final model, represented on a "virtual" map on a color scale. Again, spots are represented as circles located on the virtual map by their *x*–*y* coordinates identified by PDQuest analysis: increasing red tones correspond to increasing positive coefficients (spots upregulated after the treatment), while increasing blue tones correspond to increasing negative coefficients (spots downregulated after the treatment).

According to the results reported in Table 2, the performances of the model are very good both in fitting and in prediction, allowing us to obtain similar RMSE values for calibration and prediction (smaller than 0.05). The results are confirmed also by Fig. 4c, showing the measured versus the calculated response in fitting and prediction: the graph shows in fact very small variations along the *y*-axis.

PLS-DA on PDQuest spots: comparison with the BE-PLS-DA model

As for the nuclei dataset, the results obtained by BE-PLS-DA were compared with those obtained by the classic differential analysis by PDQuest. PLS-DA was then applied to the dataset described by the 48 spots identified as relevant by PDQuest analysis. Table 1 reports the percentage of explained and cumulative explained variance on *Y* and *X* variables, calculated for the first eight LVs. The first LV appears to be the significant one, explaining about 97% of variance on the *Y* variable and more than 60% on the *X* block. The first LV was then selected as relevant and was included in the final model.

Figure 5b shows the score plot and the loading plot of the first two LVs; the score plot shows the separation of the samples in the two groups along the first LV: control samples at negative values (circles) and treated samples at positive ones (squares). The loading plot identifies 21 spots at negative loadings on LV_1 (downregulated after the treatment) and 27 at positive loadings on the same LV (upregulated after the treatment).

Figure 3d shows the coefficients of the original variables calculated for the first LV included in the model: as for the previous case, the coefficients are represented on a virtual map on a color scale from dark blue (large negative coefficients) to dark red (large positive values). Variables with large negative coefficients correspond to spots downregulated after the treatment, while those characterized by positive coefficients are upregulated after the treatment.

The performances of the model appear quite good both in fitting and in prediction, according to the R^2 values (Table 2); however, the model performs worse than the one built with variable selection, particularly in regard to the prediction ability. This is particularly visible when looking at the RMSE values in fitting and prediction, showing that the error produced by the model based on PDQuest analysis is about 5 times larger than that produced by the model based on variable selection. The model based on multivariate analysis is thus more reliable above all for what regards prediction. This is also confirmed by the representation of calculated versus measured responses in fitting and prediction (Fig. 4d), showing large variations along the *y*-axis with respect to BE-PLS-DA model (Fig. 4c).

Table 4 reports the spots identified as relevant by BE-PLS-DA and PLS-DA based on PDQuest analysis. BE-PLS-DA identifies 89.58% of the spots identified by PDQuest (43 of 48). Only 10.42% (five of 48) are identified by PDQuest only, while BE-PLS-DA identifies 103 spots more. Therefore, the BE-PLS-DA procedure does not identify only a very small number of spots identified by PDQuest. The spots are listed according to their decreasing

Table 3 Nuclei dataset: BE-PLS-DA and PDQuest PLS-DA coefficients of the models containing the first LV and the corresponding significance level of each variable

Downregulated					Upregulated				
SSP	BE		PDQuest		SSP	BE		PDQuest	
	Coefficient	Significance	Coefficient	Significance		Coefficient	Significance	Coefficient	Significance
SSP7210	−0.0203	3.72E-14			SSP1101	0.0201	2.78E-14	0.0314	9.18E-12
SSP6509	−0.0198	5.94E-14			SSP7006	0.0201	6.46E-15		
SSP8401	−0.0197	3.76E-15	−0.0308	3.20E-12	SSP7103	0.0200	9.33E-16	0.0313	1.61E-12
SSP3207	−0.0194	4.02E-16	−0.0304	8.97E-13	SSP107	0.0196	3.73E-14	0.0307	1.01E-11
SSP608	−0.0194	1.28E-13			SSP8105	0.0195	6.24E-13		
SSP6107	−0.0192	1.48E-14	−0.0300	5.02E-12	SSP1403	0.0195	2.90E-16		
SSP7705	−0.0188	9.18E-14			SSP2902	0.0195	2.73E-13		
SSP8205	−0.0188	1.93E-12			SSP1104	0.0195	2.63E-14	0.0304	5.91E-12
SSP303	−0.0186	5.92E-12			SSP2606	0.0193	3.50E-12	0.0302	7.61E-11
SSP7211	−0.0186	2.92E-15	−0.0290	1.19E-12	SSP106	0.0193	5.76E-15	0.0302	2.91E-12
SSP6708	−0.0183	1.79E-11	−0.0287	2.19E-10	SSP1006	0.0193	3.41E-15		
SSP3510	−0.0183	2.73E-13			SSP7408	0.0192	1.86E-13		
SSP6804	−0.0181	2.05E-11	−0.0283	1.43E-10	SSP3202	0.0192	2.97E-15	0.0300	9.77E-13
SSP6808	−0.0180	5.33E-11	−0.0281	1.31E-10	SSP3005	0.0191	1.71E-14	0.0298	2.54E-12
SSP7301	−0.0177	2.82E-10			SSP4611	0.0189	2.44E-14	0.0295	8.94E-12
SSP6310	−0.0177	2.40E-10			SSP2710	0.0189	1.89E-12		
SSP3406	−0.0173	5.36E-11			SSP7311	0.0189	1.15E-11		
SSP6210	−0.0171	1.38E-12			SSP4112	0.0188	9.31E-13	0.0294	3.32E-11
SSP3508	−0.0171	9.99E-12	−0.0267	6.85E-13	SSP3106	0.0188	3.09E-13		
SSP5305	−0.0170	5.06E-12			SSP8110	0.0187	3.24E-12		
SSP8003	−0.0166	6.13E-11			SSP5504	0.0186	6.00E-13		
SSP7309	−0.0163	7.17E-10			SSP2506	0.0186	3.46E-12		
SSP10	−0.0161	6.39E-11			SSP2608	0.0183	7.49E-11		
SSP3105	−0.0159	3.46E-08			SSP6211	0.0182	1.79E-13	0.0285	6.81E-14
SSP3307	−0.0150	1.30E-07			SSP3208	0.0182	1.58E-12	0.0284	7.26E-13
SSP6901	−0.0146	1.51E-09	−0.0228	1.20E-09	SSP4604	0.0180	1.10E-11		
SSP3305	−0.0142	5.44E-08			SSP2201	0.0180	3.52E-10		
SSP1706			−0.0222	7.20E-08	SSP5004	0.0176	4.82E-12		
SSP2007			−0.0217	5.22E-08	SSP9002	0.0176	1.21E-10	0.0275	9.11E-10
SSP2105			−0.0224	1.91E-07	SSP8203	0.0174	5.41E-10		
SSP2705			−0.0196	9.49E-07	SSP6103	0.0174	1.22E-12		
SSP2712			−0.0207	1.62E-07	SSP8304	0.0174	7.78E-13		
SSP3704			−0.0176	1.22E-06	SSP7618	0.0172	3.74E-08	0.0268	6.10E-08
SSP4007			−0.0229	1.44E-06	SSP8412	0.0170	1.77E-08	0.0265	3.69E-08
SSP4505			−0.0264	3.41E-10	SSP1607			0.0230	1.66E-06
SSP6413			−0.0256	3.85E-10	SSP2207			0.0203	4.72E-07
SSP6801			−0.0202	6.20E-08	SSP4203			0.0280	5.32E-10
SSP8414			−0.0265	9.47E-09	SSP5303			0.0279	3.60E-10
SSP8504			−0.0239	1.51E-08	SSP5601			0.0212	4.51E-06
SSP8806			−0.0183	7.34E-07	SSP606			0.0221	1.69E-07
					SSP702			0.0218	7.15E-08
					SSP7508			0.0215	1.91E-06
					SSP8503			0.0259	9.19E-09

SSP standard spot

positive or negative coefficient. Table 4 also reports the PLS-DA coefficients for the model built on the spots identified as significant by PDQuest, together with the significance levels for each variable in each model. Again, the significance levels were calculated by a *t* test on each coefficient in cross-validation. Also in this case, the coefficients calculated by the BE-PLS-DA model are all significant and their values are similar to those provided by PLS-DA on the spots identified by PDQuest. However, the BE-PLS-DA procedure includes more variables that show a high relevance.

Mass spectrometry

The study was further deepened at the biological level by identifying some of regulated spots by mass spectrometry analysis (Table 5). The regulated spots subjected to the identification were those detected by both analyses (PDQuest and BE-PLS-DA) and by only the BE-PLS-DA procedure. The proteins identified help us to increase our understanding of the RC307 mechanism of action. For example, it is interesting to note that RC307 enhances the level of heat shock 70-kDa protein 5 (HSPA5) (upregulated 2 fold). This finding is in agreement with that of a previously reported study regarding the ability of valproic acid to inhibit HDAC and induce HSPA5 protein chaperone expression [31]. Another heat shock protein that we found regulated by RC307 was heat shock 90-kDa protein 1, beta (downregulated 4.5 fold). Interestingly, several hsp90 inhibitors have been developed against the different type of cancer. Recently, inhibition of hsp90 was obtained in the HCT116 cell line by using resorcinylic pyrazole/isoxazole amide analogues, resulting in HCT116 growth inhibition [32]. We also found a regulation of peroxiredoxin I

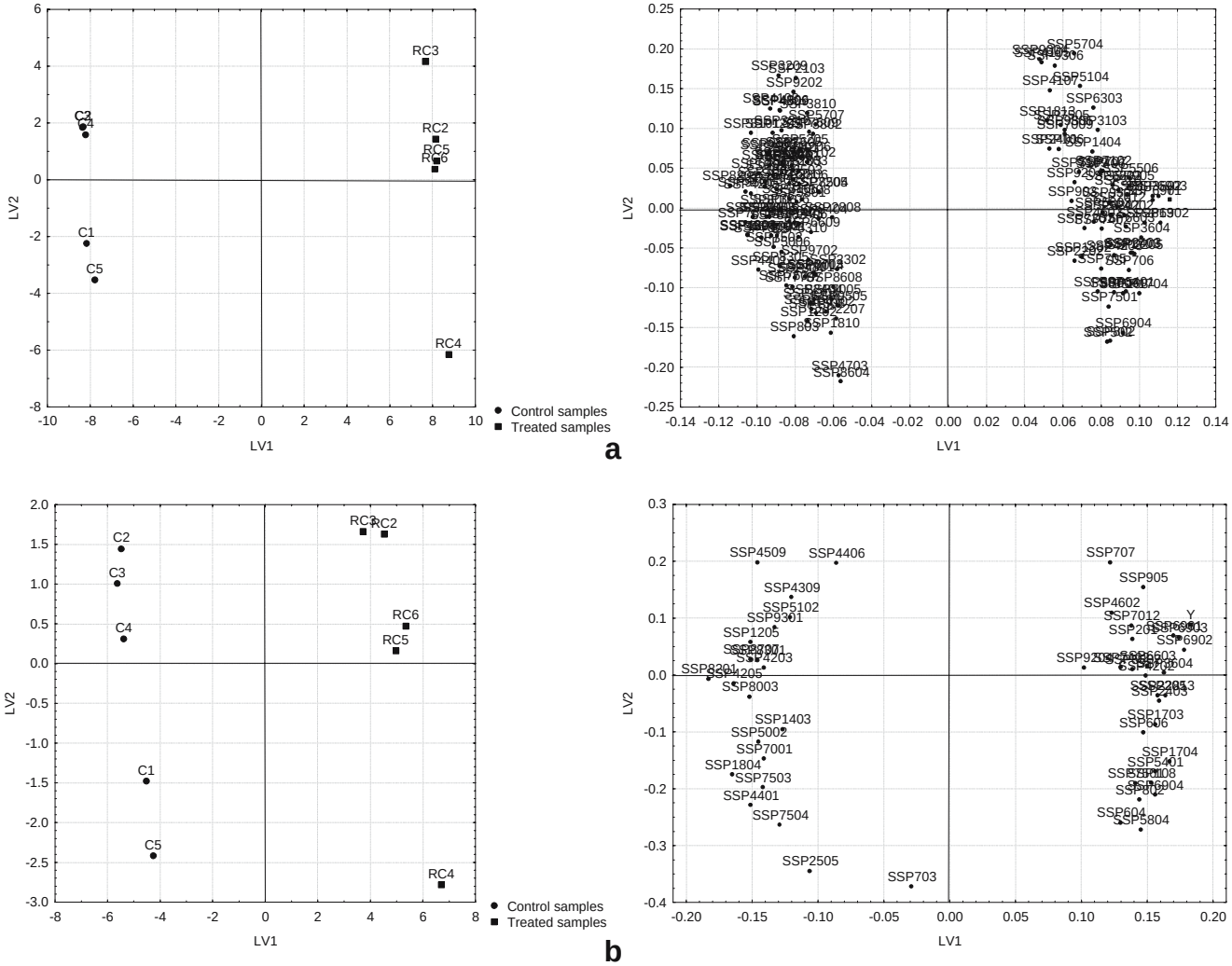

Fig. 5 Lysates dataset: score plot and loading plot (**a**) of the first two latent variables for the final PLS-DA model with variable selection; score plot and loading plot (**b**) of the first two latent variables for the PLS-DA model containing the spots identified by PDQuest. For the score plots control samples are represented as *circles* and treated samples as *squares*; for the loading plots the *X* variables are represented as *circles* and the *Y* variables as *squares*

Table 4 Lysates dataset: BE-PLS-DA and PDQuest PLS-DA coefficients of the models containing the first LV and the corresponding significance level of each variable

Downregulated					Upregulated				
SSP	BE		PDQuest		SSP	BE		PDQuest	
	Coefficient	Significance	Coefficient	Significance		Coefficient	Significance	Coefficient	Significance
SSP8201	−0.0132	8.63E-14	−0.0336	2.08E-12	SSP6902	0.0128	9.89E-12	0.0327	1.26E-11
SSP7602	−0.0123	2.66E-12			SSP6903	0.0128	2.59E-13	0.0326	1.75E-12
SSP4308	−0.0123	3.09E-11			SSP3502	0.0124	5.95E-13		
SSP1804	−0.0122	5.88E-12	−0.0313	8.19E-12	SSP6901	0.0123	1.34E-12	0.0315	3.65E-12
SSP9501	−0.0122	4.50E-13			SSP2813	0.0119	5.27E-11	0.0304	6.57E-11
SSP4205	−0.012	9.13E-12	−0.0306	8.60E-11	SSP3604	0.0117	7.05E-11	0.0298	4.85E-11
SSP5204	−0.0119	1.10E-10			SSP1704	0.0115	5.12E-10	0.0293	5.24E-10
SSP8101	−0.0118	1.47E-13			SSP2205	0.0112	1.48E-10	0.0286	5.93E-10
SSP4401	−0.0116	2.45E-10	−0.0296	2.23E-09	SSP2403	0.0111	3.07E-11	0.0285	3.72E-11
SSP6401	−0.0115	7.67E-11			SSP905	0.0111	2.14E-10	0.0283	1.51E-09
SSP7305	−0.0114	6.26E-10			SSP1703	0.011	4.21E-10	0.0282	1.75E-10
SSP7204	−0.0111	2.59E-09			SSP5506	0.011	7.64E-08		
SSP3002	−0.0111	4.79E-09			SSP1	0.0109	1.20E-08		
SSP8003	−0.011	2.51E-10	−0.0282	7.23E-10	SSP706	0.0109	3.63E-11		
SSP8006	−0.011	2.56E-08			SSP4202	0.0107	4.21E-11	0.0273	2.30E-10
SSP8005	−0.0109	2.47E-08			SSP6603	0.0107	1.45E-08	0.0273	9.43E-09
SSP5002	−0.0108	4.15E-09	−0.0277	2.33E-09	SSP5401	0.0106	1.92E-08	0.0272	4.39E-09
SSP9402	−0.0108	1.76E-08			SSP108	0.0105	6.23E-08	0.0267	9.45E-09
SSP8707	−0.0108	9.47E-09	−0.0276	1.95E-08	SSP6904	0.0104	3.75E-07	0.0265	1.36E-06
SSP9303	−0.0107	1.26E-10			SSP201	0.0104	1.39E-08	0.0265	5.87E-09
SSP4108	−0.0106	1.85E-09			SSP7012	0.0102	9.64E-09	0.0261	1.42E-08
SSP7503	−0.0106	2.75E-08	−0.0271	9.68E-08	SSP5807	0.01	9.62E-08	0.0255	5.12E-08
SSP8301	−0.0105	5.28E-11	−0.0269	6.72E-11	SSP606	0.0099	3.55E-08	0.0254	6.00E-09
SSP1205	−0.0105	1.91E-10	−0.0269	1.36E-11	SSP6802	0.0099	4.51E-09		
SSP7001	−0.0104	2.63E-09	−0.0266	1.13E-09	SSP4208	0.0099	6.24E-07		
SSP8403	−0.0104	5.75E-09			SSP9304	0.0098	4.21E-09		
SSP5003	−0.0104	5.14E-08			SSP802	0.0096	6.14E-08	0.0245	1.12E-08
SSP8305	−0.0103	3.08E-10			SSP7501	0.0095	5.98E-08	0.0243	1.80E-08
SSP3206	−0.0102	1.41E-09			SSP7102	0.0095	2.37E-07		
SSP5006	−0.0101	4.80E-08			SSP502	0.0094	2.14E-07		
SSP3209	−0.0101	9.26E-10			SSP707	0.0094	4.73E-08	0.0240	3.20E-08
SSP4806	−0.0101	8.54E-08			SSP5803	0.0094	7.68E-08		
SSP4509	−0.0101	1.00E-07	−0.0257	5.62E-08	SSP406	0.0093	7.27E-08		
SSP5306	−0.00998	1.79E-09			SSP705	0.0092	5.31E-09		
SSP3210	−0.00998	2.79E-08			SSP7402	0.0092	2.03E-07	0.0236	3.05E-07
SSP5103	−0.00995	1.28E-09			SSP3103	0.0092	6.69E-07		
SSP7504	−0.00992	7.79E-09	−0.0254	2.13E-08	SSP6303	0.009	2.55E-06		
SSP4203	−0.00992	6.02E-08	−0.0254	7.63E-08	SSP607	0.0089	1.72E-06		
SSP7601	−0.00964	4.14E-06			SSP1404	0.0089	3.87E-07		
SSP8002	−0.00963	1.14E-08			SSP4602	0.0088	3.32E-06	0.0226	1.78E-06
SSP7105	−0.00961	5.30E-07			SSP7301	0.0082	3.26E-05		
SSP803	−0.00956	4.34E-08			SSP5104	0.0082	1.62E-05		
SSP7707	−0.0095	2.85E-09			SSP1302	0.0081	3.80E-05		
SSP9301	−0.00949	5.87E-09	−0.0243	2.01E-09	SSP5704	0.0079	7.75E-05		
SSP8103	−0.00938	1.10E-07			SSP7109	0.0079	8.79E-07		
SSP8601	−0.00936	2.73E-05			SSP9204	0.0077	8.93E-05	0.0196	8.43E-05
SSP1403	−0.00936	1.38E-07	−0.0239	9.11E-08	SSP2202	0.0076	4.98E-05		
SSP2304	−0.00932	4.18E-07			SSP903	0.0076	4.18E-06		
SSP7605	−0.0093	6.15E-08			SSP7009	0.0072	2.13E-04		
SSP6503	−0.00928	7.60E-08			SSP9305	0.0072	1.15E-04		
SSP9202	−0.00917	1.17E-07			SSP7505	0.0069	1.35E-04		
SSP6310	−0.00913	7.34E-07			SSP106	0.0068	3.36E-04		

Table 4 (continued)

Downregulated					Upregulated				
SSP	BE		PDQuest		SSP	BE		PDQuest	
	Coefficient	Significance	Coefficient	Significance		Coefficient	Significance	Coefficient	Significance
SSP5301	−0.00911	1.89E-06			SSP9306	0.0067	2.42E-04		
SSP2103	−0.00907	1.77E-06			SSP4107	0.0064	5.76E-04		
SSP2806	−0.00906	3.59E-05			SSP2401	0.0063	3.68E-04		
SSP7406	−0.00902	6.00E-07			SSP1813	0.0062	7.35E-04		
SSP2303	−0.00897	4.29E-06			SSP4105	0.0059	1.26E-03		
SSP5205	−0.00888	1.45E-05			SSP9004	0.0058	6.34E-04		
SSP6308	−0.00886	6.07E-07			SSP5804			0.0239	4.91E-07
SSP4206	−0.00881	1.67E-08			SSP604			0.0219	1.30E-05
SSP4801	−0.00874	1.91E-07							
SSP1202	−0.0087	1.35E-06							
SSP9702	−0.00849	2.61E-06							
SSP4304	−0.00848	3.78E-05							
SSP2505	−0.00848	4.04E-07	−0.0217	1.16E-06					
SSP5102	−0.00847	2.53E-07	−0.0216	1.95E-07					
SSP3810	−0.00843	2.13E-06							
SSP7609	−0.00835	7.35E-06							
SSP6609	−0.00832	3.65E-05							
SSP2404	−0.00829	1.16E-05							
SSP8703	−0.00816	1.59E-05							
SSP1508	−0.00813	1.44E-05							
SSP7014	−0.00809	1.01E-05							
SSP3802	−0.00809	6.77E-05							
SSP5707	−0.00794	3.99E-06							
SSP8404	−0.00788	2.31E-06							
SSP7205	−0.00782	1.42E-04							
SSP8504	−0.00781	2.62E-06							
SSP4702	−0.00751	1.65E-04							
SSP1810	−0.00728	2.34E-05							
SSP7005	−0.00717	4.56E-06							
SSP2808	−0.00712	1.85E-03							
SSP8608	−0.00699	2.14E-04							
SSP2207	−0.00695	9.47E-05							
SSP4703	−0.00686	1.00E-03							
SSP2302	−0.00681	1.62E-04							
SSP9505	−0.00679	8.18E-04							
SSP8604	−0.00674	1.60E-04							
SSP703			−0.0084	9.03E-02					
SSP4309			−0.0213	2.29E-07					
SSP4406			−0.0145	4.72E-03					

Table 5 Identification of the differentially expressed proteins after treatment with RC307 in the HCT116 cell line

Protein	SSP	Extract	HUGO gene name	NCBI accession no.	No. of peptides	MASCOT score	Fold of variation in treated cells
Heat shock 70-kDa protein 5 (grp78)	702	Nuclear	HSPA5	gi|386758	1	112	Up 2,09
Heat shock 90-kDa protein 1, beta	1804	Total lysate	HSP90AB1	gi|20149594	19	756	Down 4,52
Peroxiredoxin 1	7103	Nuclear	PRDX1	gi|4505591	13	464	Up 2,58
Stathmin 1	4205	Total lysate	STMN1	gi|5031851	20	685	Down 2,29

(PRDX1) (upregulated 2.6 fold). PRDX1 is a scavenger of reactive oxygen species, and interacts with a region of the oncogene c-Myc that is essential for neoplastic transformation. It has been reported that this results in the suppression of some c-Myc functions, implying a tumor suppressor role for PRDX1 [33]. Moreover, it has also been demonstrated that the HDAC inhibitor FK228 activates the tumor suppressor PRDX1, inducing apoptosis in esophageal cancer cells [34]. Interestingly, we also found that the RC307 HDAC inhibitor downregulates the stathmin 1 (STMN1) protein in the treated HCT116 cells. STMN1 (downregulated 2.6 fold) is an 18-kDa ubiquitously expressed cytosolic protein that destabilizes microtubules through tubulin sequestration and/or promotion of microtubule catastrophe [35]. Moreover, STMN1 is transcriptionally repressed by functional p53 and downregulation of STMN1 is associated with G2/M cell cycle arrest [36]. Our data suggest that STMN1 downregulation plays a role in the mechanism of action of the drug tested, involving a cell cycle block at G2/M, as reported for others HDAC inhibitors (such as trichostatin A). Taking into account all the proteins identified, we could confirm the RC307 hypothesized antineoplastic effect on the HCT116 cell line.

Conclusions

PLS-DA was applied coupled with variable selection procedures to identify the differentially expressed proteins in two case-studies: nuclei and lysates from the human colon cancer cell line HCT116 treated by the HDAC inhibitor RC307. The results obtained were compared with those provided by a classic PDQuest differential analysis based on individual Student t tests (α=0.05).

The multivariate procedure applied with variable selection provided the best results regarding both the identification of a higher number of significant spots and the achievement of a model characterized by the best predictive ability. The results proved that PDQuest differential analysis identifies a subset of significant spots but this subset does not take into account the multivariate structure of the data and the relationships among the spot volumes. Multivariate tools provide better results above all regarding the predictive performances since they are able to account for the relationships existing among the variables (spots) and their correlation.

The study was further deepened by the identification of some of the most relevant spots by MALDI-TOF-TOF and RP-HPLC-ESI MS/MS. As most relevant, we chose the spots with the largest PLS-DA coefficients (the most upregulated or downregulated after the treatment). The preliminary data obtained are encouraging and some hypotheses were also drawn on the biological role played by these proteins. The modulation of the proteins identified reported is consistent with the observation that the new HDAC inhibitor studied has an antitumoral effect. Work is in progress to extend the present data along the following lines: (1) completing the mass spectrometry protein identification; (2) following the kinetics of drug exposure, in order to analyze the timing of the upregulation and downregulation profiles.

Acknowledgements The authors thank A. Vindigni and L. Zolla for mass spectrometry experiments. This work was supported in part by a grant from Fondazione Cariplo to P.G.R.

References

1. Anderson NL, Hofmann JP, Gemmell A, Taylor J (1984) Clin Chem 30:2031–2036
2. Tarroux P, Vincens P, Rabilloud T (1987) Electrophoresis 8:187–199
3. Gadea I, Ayala G, Diago MT, Cunat A, Garcia de Lomas J (2000) Clin Diagn Lab Immunol 7:549–552
4. Kovarova H, Hajduch M, Korinkova G, Halada P, Krupickova S, Gouldsworthy A, Zhelev N, Strnad M (2000) Electrophoresis 21:3757–3764
5. Iwadate Y, Sakaida T, Hiwasa T, Nagai Y, Ishikura H, Takiguchi M, Yamaura A (2004) Cancer Res 64:2496–2501
6. Amin RA, Vickers AE, Sistare F, Thompson KL, Roman RJ, Lawton M, Kramer J, Hamadeh HK, Collins J, Grissom S, Bennett L, Tucker CJ, Wild S, Kind C, Oreffo V, Davis JW, Curtiss S, Naciff JM, Cunningham M, Tennant R, Stevens J, Car B, Bertram TA, Afsharil CA (2004) Environ Health Perspect 112:465–479
7. Heijne WHM, Stierum RH, Slijper M, van Bladeren PJ, van Ommen B (2003) Biochem Pharmacol 65:857–875
8. Verhoeckx KCM, Bijlsma S, de Groene EM, Witkamp RF, van der Greef J, Rodenburg RJT (2004) Proteomics 4:1014–1028
9. Magdic D, Horvat D, Jurkovic Z, Sudar R, Kurtanjek K (2002) Food Technol Biotechnol 40:331–341
10. Marengo E, Robotti E, Righetti PG, Campostrini N, Pascali J, Ponzoni M (2004) Clin Chim Acta 345:55–67
11. Marengo E, Robotti E, Cecconi D, Scarpa A, Righetti PG (2004) Anal Bioanal Chem 379(7–8):992–1003
12. Marengo E, Robotti E, Antonucci F, Cecconi D, Campostrini N, Righetti PG (2005) Proteomics 5:654–666
13. Marengo E, Robotti E, Bobba M, Liparota MC, Antonucci F, Rustichelli C, Zamò A, Chilosi M, Hamdan M, Righetti PG (2006) Electrophoresis 27:484–494
14. Lilley KS, Cupree P (2006) J Exp Bot 57:1493–1499
15. Pir P, Kirdar B, Hayes A, Onsan ZY, Ulgen KO, Oliver SG (2006) BMC Bioinformatics 7:203
16. Verhoeckx KCM, Gaspari M, Bijlsma S, van der Greef J, Witkamp RF, Doornbos RP, Rodenburg RJT (2005) J Proteome Res 4:2015–2023
17. Jessen F, Lametsch R, Bendixen E, Kjaersgard IVH, Jorgensen BM (2002) Proteomics 2:32–35
18. Kleno TG, Leonardsen LR, Kjeldal HO, Laursen SM, Jensen ON, Baunsgaard D (2004) Proteomics 4:868–880
19. Kjaersgard IVH, Norrelykke MR, Jessen F (2006) Proteomics 6:1606–1618
20. Karp NA, Griffin JL, Lilley KS (2005) Proteomics 5:81–90
21. Norden B, Broberg P, Lindberg C, Plymoth A (2005) Chem Biodivers 2:1487–1494

22. Garcia JS, Silva GA, Arruda MAZ, Poppi RJ (2007) X-Ray Spectrom 36:122–129
23. Marks PA, Richon VM, Rifkind RA (2000) J Natl Cancer Inst 92:1210–1216
24. Cecconi D, Donadelli M, Rinalducci S, Zolla L, Scarpa A, Palmieri M, Righetti PG (2007) Proteomics 7:1644–1653
25. Cecconi D, Scarpa A, Donadelli M, Palmieri M, Hamdan M, Astner H, Righetti PG (2003) Electrophoresis 24:1871–1878
26. Massart DL, Vandeginste BGM, Deming SM, Michotte Y, Kaufman L (1988) Chemometrics: a textbook. Elsevier, Amsterdam
27. Vandeginste BGM, Massart DL, Buydens LMC, De Jong S, Lewi PJ, Smeyers-Verbeke J (1998) Handbook of chemometrics and qualimetrics: part B. Elsevier, Amsterdam
28. Eisenbeis RA (ed) (1972) Discriminant analysis and classification procedures: theory and applications. Lexington, Lexington
29. Martens H, Naes T (1989) Multivariate calibration. Wiley, London
30. Kleinbaum D, Kupper L, Muller K (1988) Applied regression analysis and other multivariate methods, 2nd edn. PWS-Kent, Boston
31. Shi Y, Gerritsma D, Bowes AJ, Capretta A, Werstuck GH (2007) Bioorg Med Chem Lett 17(16):4491–4494
32. Sharp SY, Prodromou C, Boxall K, Powers MV, Holmes JL et al (2007) Mol Cancer Ther 6(4):1198–1211
33. Egler RA, Fernandes E, Rothermund K, Sereika S, de Souza-Pinto N, Jaruga P, Dizdaroglu M, Prochownik EV (2005) Oncogene 24:8038–8050
34. Hoshino I, Matsubara H, Hanari N, Mori M, Nishimori T, Yoneyama Y, Akutsu Y, Sakata H, Matsushita K, Seki N, Ochiai T (2005) Clin Cancer Res 11(21):7945–7952
35. Cassimeris L (2002) Curr Opin Cell Biol 14:18–24
36. Johnsen JI, Aurelio ON, Kwaja Z et al (2000) Int J Cancer 88:685–691

Biochimica et Biophysica Acta 1784 (2008) 1702–1710

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbapap

A proteomic approach for evaluating the cell response to a novel histone deacetylase inhibitor in colon cancer cells

Alberto Milli [a], Daniela Cecconi [a,*], Natascia Campostrini [b], Anna Maria Timperio [c], Lello Zolla [c], Sabina Carla Righetti [d], Franco Zunino [d], Paola Perego [d], Valentina Benedetti [d], Laura Gatti [d], Federico Odreman [e], Alessandro Vindigni [e], Pier Giorgio Righetti [f]

[a] Dipartimento Scientifico e Tecnologico, Laboratorio di Proteomica, Università degli Studi di Verona, Verona, Italy
[b] Dipartimento di Medicina Clinica e Sperimentale, Università degli Studi di Verona, Verona, Italy
[c] Dipartimento di Scienze Ambientali, Università degli Studi della Tuscia, Viterbo, Italy
[d] Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy
[e] Centro Internazionale di Ingegneria Genetica e Biotecnologia "ICGEB", Trieste, Italy
[f] Dipartimento di Chimica, Materiali e Ingegneria Chimica "Giulio Natta", Politecnico di Milano, Milan, Italy

ARTICLE INFO

Article history:
Received 17 March 2008
Received in revised form 17 April 2008
Accepted 24 April 2008
Available online 4 May 2008

Keywords:
Histone deacetylases
Proteomics
Colon cancer

ABSTRACT

Epigenetic inactivation of gene expression is a general phenomenon associated with malignant transformation. Recently, we have found that a novel series of histone deacetylases (HDAC) inhibitors exhibit a broad-spectrum inhibition profile characterized by a marked effect on acetylation of histone and non-histone proteins. RC307, a representative compound of this series, caused a growth-inhibitory effect in colon carcinoma cells HCT116 associated with G2 accumulation and induction of apoptosis. The present study was designed to investigate the effect of RC307 on protein expressions in the HCT116 cells following treatment with cytotoxic drug concentrations. HCT116 cells were cultured in the absence or presence of RC307 and total cell lysates, as well as nuclear proteins, were extracted. The protein samples were then subjected to two-dimensional polyacrylamide gel electrophoresis, and the 2D gel images were compared to discover the protein changes caused by RC307 treatment. A total of 48 and 46 different spots were found to be modulated by RC307 in total lysates and nuclear proteome of HCT116 cell line. The modulated proteins were identified by tandem mass spectrometry. We found that RC307 exposure modulates proteins that are involved in proliferation, cell cycle regulation, apoptosis, gene expression, as well as chromatin and cytoskeleton organization.

1. Introduction

The acetylation state of several cellular proteins (including histones and transcription factors) is controlled by the action of histone acetyl transferases and histone deacetylases (HDAC) [1]. The post-translational modification of histones has emerged as a central mechanism of transcriptional control, because hypoacetylation results in transcriptional repression as a consequence of tight packing of DNA into the nucleosome. Modulation of histone acetylation and chromatin relaxation affect diverse cellular functions including control of cell growth, differentiation and apoptosis [2]. In addition to modulation of histone acetylation, HDACs are also involved in lysine acetylation state of non-histone proteins implicated in regulatory processes (e.g., transcriptional factors and tubulin) [1]. Epigenetic inactivation of gene expression is a general phenomenon associated with malignant transformation [3]. Indeed, the expression of several regulatory genes, including tumour suppressor genes, differentiation genes and DNA repair genes may be repressed during tumour transformation and progression. For these reasons, modulation of epigenetic gene repression has been proposed as an attractive approach to control tumour growth, and HDACs are recognized as potential targets of this approach [4,5]. Class I and class II HDACs are zinc-dependent metalloenzymes that catalyze the hydrolysis of acetylated lysine residues. Most HDAC inhibitors have a zinc binding group, such as hydroxamic acid. Therefore, these HDAC inhibitors are not isoenzyme-specific and could be considered as pan-HDAC inhibitors. We have already explored the effects of the hydroxamic acid-based HDAC inhibitor trichostatin A (TSA) on pancreatic exocrine [6,7] and endocrine [8] cancer cell lines, showing that drug-induced inhibition of proliferation was associated with cell cycle arrest and apoptosis and TSA produced a profound change in the global proteomic profile. In spite of much effort, the design of selective inhibitors remains difficult and only few compounds exhibited some degree of specificity [9,10]. It remains to be defined if isoenzyme-specific HDAC inhibitors provide advantages over aspecific pan-HDAC inhibitors. Recently, we have found that a novel series of HDAC

* Corresponding author. Dipartimento Scientifico e Tecnologico, Laboratorio di Proteomica, Università degli Studi di Verona, Strada le Grazie 15, 37134, Verona, Italy. Fax: +39 045 8027929.
E-mail address: daniela.cecconi@univr.it (D. Cecconi).

1570-9639/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbapap.2008.04.022

inhibitors containing hydroxamic acid exhibit a broad-spectrum inhibition profile characterized by a marked effect on acetylation of histone and non-histone proteins (manuscript in preparation). We have found that RC307, a representative compound of this series, was able to induce apoptosis in ovarian carcinoma cells (manuscript in preparation). In the human colon carcinoma cells HCT116, known to be a cell system responsive to HDAC inhibitors [11], RC307 exhibited an antiproliferative effect that was associated with G2-phase accumulation but only a moderate induction of apoptosis. Based on these observations, the present study was designed to investigate the pattern of protein expression in tumour cells following treatment with antiproliferative concentrations of RC307. In addition, our proteomic findings may elucidate the critical protein expression alterations responsible for the activity of this novel HDAC inhibitor. Understanding the mechanism of action of RC307 will finally also allow an improvement in knowledge of the epigenetic events involved in colon cancer development.

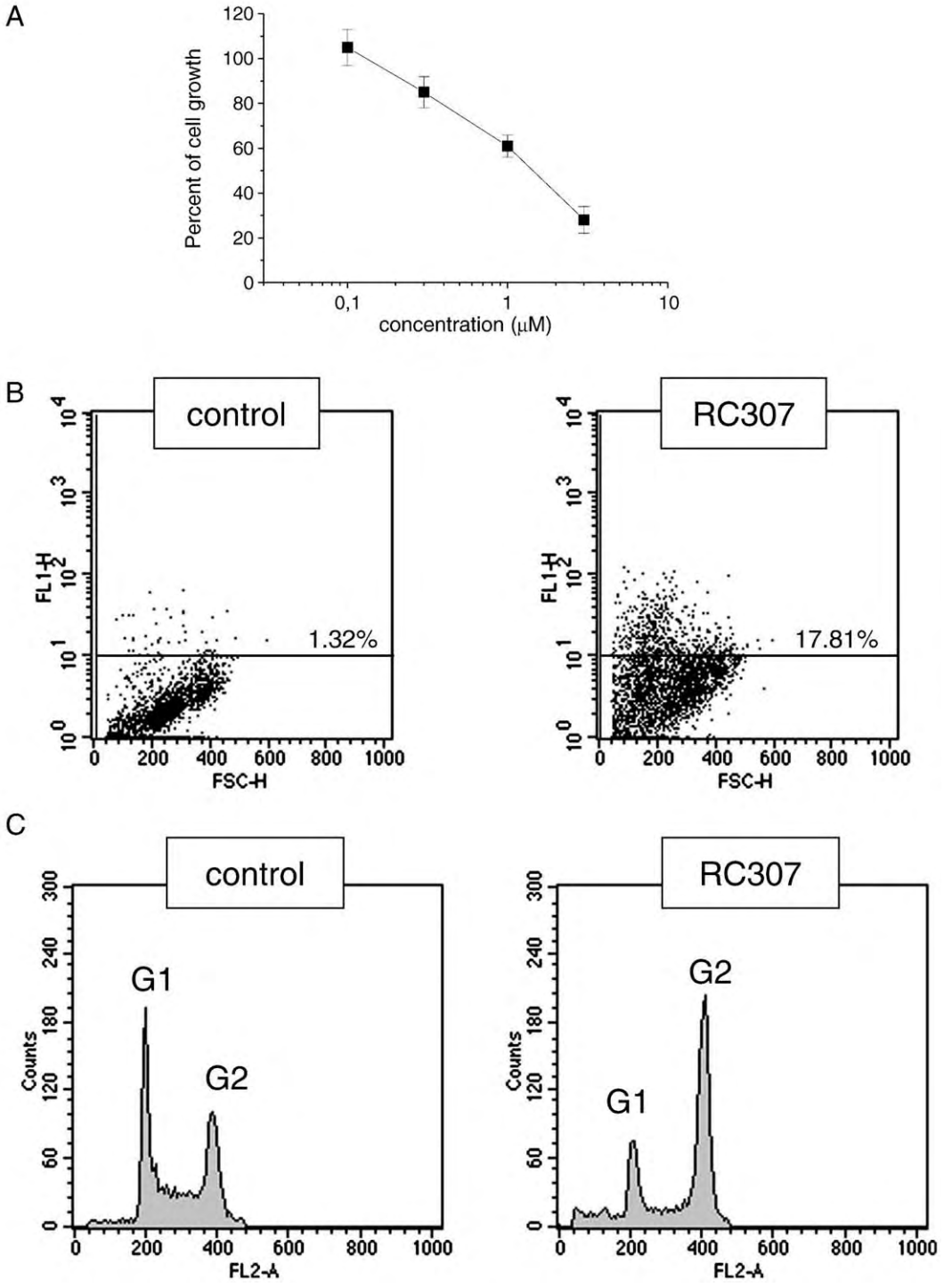

Fig. 1. (A) Sensitivity of HCT116 cells to RC307. Sensitivity was assessed by growth inhibition assay with a 72 h exposure and cells were counted at the end of treatment. Values are the mean (±SD) of six independent experiments. (B) Analysis of apoptosis in HCT116 cells exposed to RC307. Apoptosis was measured by TUNEL assay after 72 h exposure to a concentration of RC307 corresponding to the IC_{80} (2.5 μM). Representative dot plots showing viable cells (FL1<10^1) and apoptotic cells (FL1>10^1) are presented. (C) Cell cycle distribution in HCT116 cells exposed to RC307. Cell cycle distribution was analyzed 24 h after exposure to RC307 (concentration corresponding to IC_{80}) in propidium iodide-stained cells. Histograms from a representative cytofluorimetric analysis are shown.

2. Materials and methods

2.1. Cell culture and drugs

The human colon cancer HCT116 cell line was grown and maintained as monolayers in RPMI-1640 medium supplemented with 10% fetal calf serum (Invitrogen, Carlsbad, CA). RC307 was dissolved in dimethylsulfoxide and diluted in water.

2.2. Growth-inhibition assays

Cell sensitivity to the HDAC inhibitor RC307 was assessed by growth-inhibition assay. Exponentially growing cells were harvested, seeded into 6-well plates and 24 h later cells were exposed to different concentrations of RC307 (range: 0.1–3 μM) or to solvent for 72 h. At the end of treatment, culture medium was removed and adherent cells were harvested using trypsin and counted with a cell counter (Coulter Electronics, Luton, UK). IC_{50} is defined as the concentration causing a 50% inhibition of cell growth as compared with control.

2.3. Apoptosis and cell cycle analysis

Exponentially growing cells were seeded in 75 cm^2 flasks and, 24 h later, they were exposed to a drug concentration corresponding to IC_{80} (2.5 μM) for 24, 48 or 72 h. Under these conditions, the concentration of the drug solvent (0.005%) had no effects. At the end of treatment floating and adherent cells were harvested for detection of apoptotic cells or cell cycle analysis. Apoptosis was evaluated by TUNEL (Terminal deoxynucleotidyl transferase dUTP Nick End Labelling) assay (Roche, Mannheim, Germany). After harvesting, cells were fixed in paraformaldehyde, permeabilized in a solution of 0.1% Triton X-100 in 0.1% sodium citrate, and then incubated in the TUNEL reaction for 1 h. After washing, samples were analyzed by flow cytometry using Cell Quest software (Becton Dickinson, Mountain View, CA). For cell cycle analysis, cells were fixed and stained with a propidium iodide (PI)-containing solution (30 μg/ml PI, 66 U/ml RNase A in PBS). The cell cycle perturbations were measured by using a flow cytometer (Becton Dickinson, Franklin Lakes, NJ). Samples were analyzed for DNA content and cell cycle distributions were calculated using Modfit software (Becton Dickinson).

2.4. Total lysates and nuclear protein extraction

Exponentially growing cells were seeded in 150 cm^2 flasks and, 24 h later, they were exposed to the HDAC inhibitor for 24 h (IC_{80}: 2.5 μM). Total lysates protein extraction from 10^8 untreated and RC307-treated cells was performed in a 2-D solubilising/lysing solution: 7 M urea (Sigma, Sigma-Aldrich Corporation, St. Louis, MO, USA), 2 M thiourea (Sigma), 3% CHAPS (Sigma), 20 mM Tris (Sigma), 1% pH 3–10 Ampholyte (Fluka, Buchs SG Switzerland) and 1× protease inhibitor cocktail tablet (Complete, Mini; Roche, Basel, Switzerland). The samples were then sonicated 5×30 s on ice with 1 min rest in between times and the sonicates were centrifuged for 10 min at 10,000×*g* at 4 °C to remove the nucleic acids complexed with ampholytes. Concerning

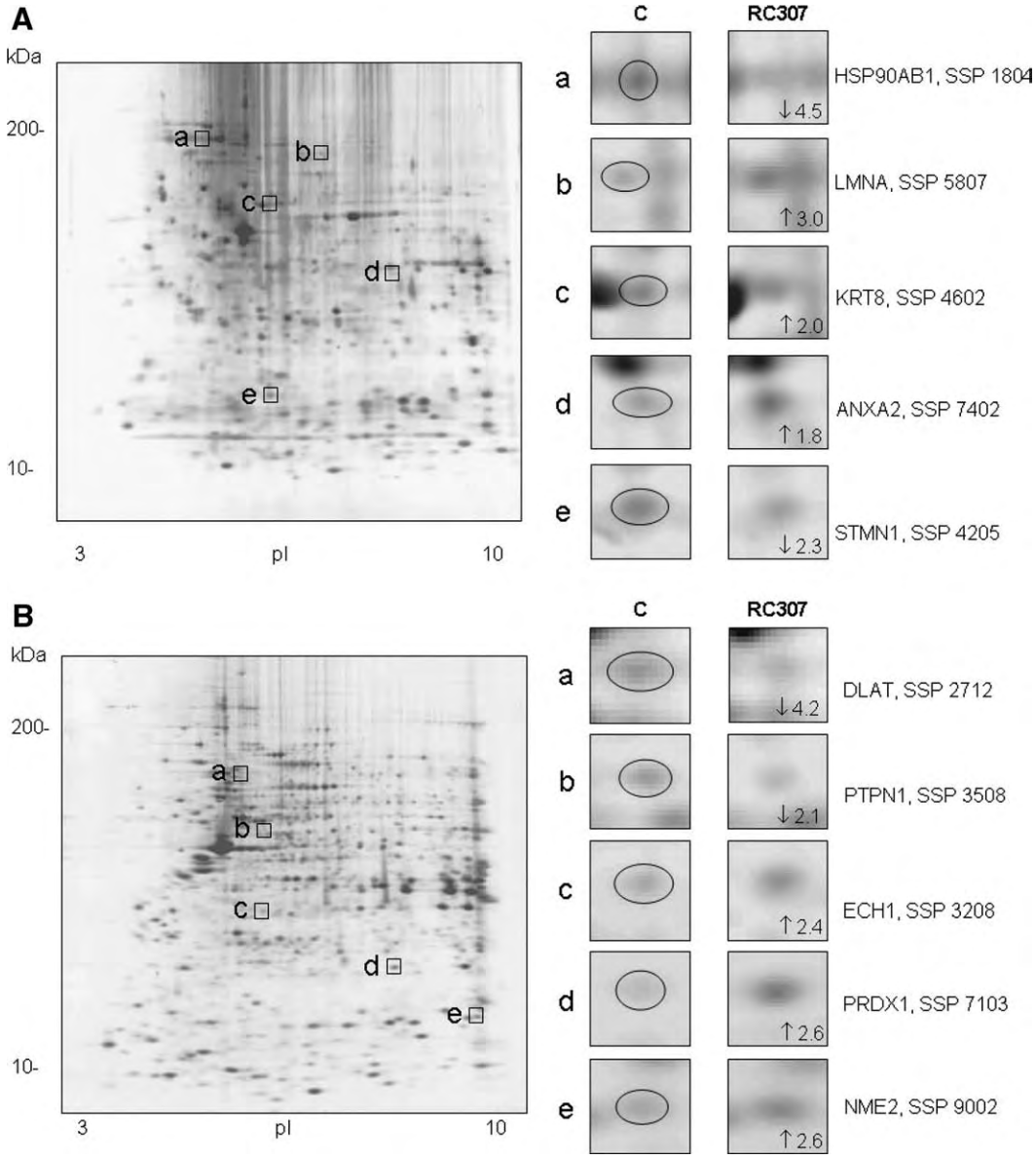

Fig. 2. High-resolution two-dimensional Gel Electrophoresis. Example of Sypro Ruby stained 2-D gels of (A) total lysates and (B) nuclear proteins extracted from HCT116 cell line, and PDQuest (version 7.3) output showing representative differentially expressed spots ($p<0.05$) between control and RC307-treated samples. For each spot, an enlarged region of the respective 2-DE map is shown which contains the referred spot (highlighted by an ellipse) and the corresponding fold of change. Each spot was identified by MS/MS analysis.

Table 1
Identified proteins that are differentially expressed at 24 h of RC307 treatment through 2-DE

Protein name	SSP	Extract	HUGO Gene Name	NCBI accession no.	No. of peptides	Mascot Score	Fold of variation in treated cells	Cellular Component GO term	Molecular Function GO term
Cell proliferation, cell cycle and apoptosis									
Adenylate kinase isoenzyme 1	9204	total lysate	AK1	gi|49456961	11	102	up 1,55	cytosol	adenylate kinase activity
Annexin V	2007	nuclear	ANXA5	gi|1421662	4	133	down 2,70	cytosol	calcium binding
Calreticulin	905	total lysate	CALR	gi|48146257	9	108	up 2,01	endoplasmic reticulum	calcium ion binding
Galectin 3	9301	total lysate	LGALS3	gi|126678	4	143	down 1,65	nucleus	immunoglobulin binding
Glutathione S-transferase pi	4202	total lysate	GSTP1	gi|47496669	7	89	up 1,61	cytosol	glutathione transferase activity
Heat shock 70 kDa protein 1 (hsp72)	1706	nuclear	HSPA1	gi|188488	1	71	down 2,78	endoplasmic reticulum	ATP binding, protein binding
Heat shock 70 kDa protein 5 (grp78)	702	nuclear	HSPA5	gi|386758	1	112	up 2,09	endoplasmic reticulum	ATP binding
Heat shock 90 kDa protein 1, beta	1804	total lysate	HSP90AB1	gi|20149594	19	756	down 4,52	cytosol	ATP binding
Inosine monophosphate dehydrogenase	4611	nuclear	IMPDH2	gi|44979607	5	244	up 2,03	cytosol	oxidoreductase activity
Parathymosin	201	total lysate	PTMS	gi|46276863	4	231	up 1,54	nucleus	protein binding
Peroxiredoxin 1	7103	nuclear	PRDX1	gi|4505591	13	464	up 2,58	cytosol	antioxidant activity
Rho GDP dissociation inhibitor	1104	nuclear	ARHGDIA	gi|3608	3	123	up 2,65	cytoskeleton	enzyme regulator activity
Septin 2	4505	nuclear	SEPT2	gi|1040689	2	104	down 2,22	cytoskeleton	pyrophosphatase activity
Stathmin 1	4205	total lysate	STMN1	gi|5031851	20	685	down 2,29	cytoskeleton	protein binding
TNF receptor-associated protein 1 (hsp75)	5804	total lysate	TRAP1	gi|1082886	4	180	up 1,96	cytosol	protein binding
Regulation of gene expression									
Alpha-enolase	6901	total lysate	ENO1	gi|39644850	9	124	up 7,18	cytosol, nucleus	transcription factor activity
Alpha-enolase	6902	total lysate	ENO1	gi|39644850	8	94	up 111,05	cytosol, nucleus	transcription factor activity
Alpha-enolase	6903	total lysate	ENO1	gi|39644850	7	85	up 14,76	cytosol, nucleus	transcription factor activity
Alpha-enolase	6904	total lysate	ENO1	gi|39644850	6	67	up 111,90	cytosol, nucleus	transcription factor activity
Alpha-enolase	6603	total lysate	ENO1	gi|39644850	10	106	up 2,17	cytosol, nucleus	transcription factor activity
Cytokine induced protein 29 kDa	4203	nuclear	CIP29	gi|32129199	2	113	up 2,51	nucleus	protein binding
Non-metastatic cells 1, protein	1101	nuclear	NME1	gi|35068	3	96	up 7,34	nucleus	transcription factor activity
Non-metastatic cells 2, protein	9002	nuclear	NME2	gi|1421609	5	288	up 2,63	nucleus	nucleoside diphosphate kinase activity
Nuclease sensitive element-binding protein 1	604	total lysate	YBX1	gi|117938841	13	138	up 1,56	nucleus	nucleic acids binding
Proliferation-associated 2G4, 38 kDa	4505	nuclear	PA2G4	gi|4099506	7	336	down 2,22	nucleus	amino peptidase activity
Profilin I	106	nuclear	PFN1	gi|999511	3	86	up 2,59	nucleus	protein binding
TBP-associated factor 15	6804	nuclear	TAF15	gi|119600533	3	96	up 2,38	nucleus	nucleic acid binding, cation binding
Signal transduction									
Annexin III	2207	nuclear	ANXA3	gi|1421662	4	243	up 2,76	cytosol	pyrophosphatase activity
Protein tyrosine phosphatase, non-receptor type 1	3508	nuclear	PTPN1	gi|809208	5	215	down 2,13	endoplasmic reticulum	protein tyrosine phosphatase activity
Chromatin and cytoskeleton organization									
Annexin II	5303	nuclear	ANXA2	gi|4757756	6	388	up 3,04	soluble fraction and plasma membrane	phospholipase inhibitor activity
Annexin II	7402	total lysate	ANXA2	gi|73909156	15	204	up 1,8	soluble fraction and plasma membrane	phospholipase inhibitor activity
Annexin II	7501	total lysate	ANXA2	gi|73909156	12	94	up 1,53	soluble fraction and plasma membrane	phospholipase inhibitor activity
Annexin IV	3202	nuclear	ANXA4	gi|189617	2	125	up 7,03	cytosol	phospholipase inhibitor activity
Beta-actin	1607	nuclear	ACTB	gi|28336	4	226	up 2,65	actin filament	protein binding
Cytokeratin 8	4602	total lysate	KRT8	gi|62913980	14	157	up 2,03	cytoskeleton	protein binding
Cytokeratin 8	3604	total lysate	KRT8	gi|62913980	17	212	up 1,94	cytoskeleton	protein binding
Cytokeratin 8	2606	nuclear	KRT8	gi|181573	23	874	up 16,74	cytoskeleton	protein binding
Cytokeratin 18	1101	nuclear	KRT18	gi|30311	5	258	up 7,34	cytoskeleton	protein binding
Histone H4	7001	total lysate	HIST4H4	gi|51317339	7	88	down 4,91	nucleus	DNA binding

(continued on next page)

Table 1 *(continued)*

Protein name	SSP	Extract	HUGO Gene Name	NCBI accession no.	No. of peptides	Mascot Score	Fold of variation in treated cells	Cellular Component GO term	Molecular Function GO term
Chromatin and cytoskeleton organization									
Lamin A/C	5804	total lysate	LMNA	gi|5031875	10	290	up 1,96	lamin filament	protein binding
Lamin A/C	5807	total lysate	LMNA	gi|125962	9	124	up 2,96	lamin filament	protein binding
RNA splicing, processing, and translation									
Cleavage and polyadenylation specific factor 6, 68 kDa	2712	nuclear	CPSF6	gi|5901928	1	97	down 4,17	nucleus	RNA binding
hnRNP D	6413	nuclear	HNRPD	gi|181914	2	64	down 2,17	nucleus	RNA and DNA binding
hnRNP D	8414	nuclear	HNRPD	gi|508270	5	153	down 3,70	nucleus	RNA and DNA binding
hnRNP H1	3508	nuclear	HNRPH1	gi|5031753	3	137	down 2,13	nucleus	RNA binding
hnRNP K	3704	nuclear	HNRPK	gi|473911	9	386	down 2,56	nucleus	RNA binding
hnRNP R	2705	nuclear	HNRPR	gi|5031755	2	128	down 2,50	nucleus	RNA binding
Mitochondrial ribosomal protein L12	4203	total lysate	MRPL12	gi|20981709	7	86	down 1,71	mitochondrion	RNA binding
Mitochondrial ribosomal protein S22	4509	total lysate	MRPS22	gi|13633893	14	225	down 1,75	mitochondrion	RNA binding
Nuclear RNA-binding protein, 54-kD	8504	nuclear	NONO	gi|543010	4	163	down 1,79	nucleus	RNA and DNA binding
Nucleolin	702	nuclear	NCL	gi|34534595	5	242	up 2,09	nucleus	RNA and DNA binding
Ribosomal protein L23a	201	total lysate	RPL23A	gi|404015	2	128	up 1,54	ribosome	RNA binding
RNA helicase 2	6808	nuclear	DHX15	gi|2696613	5	150	down 2,78	nucleus	RNA helicase activity
Splicing factor SF3a60	1607	nuclear	SF3A3	gi| 551450	9	533	up 2,65	nucleus	RNA binding
Splicing factor, arginine/serine rich 1	2403	total lysate	SFRS1	gi|5902076	5	254	up 1,73	nucleus	RNA binding
Translation initiation factor 3, p40 subunit	4406	total lysate	EIF3S3	gi|3986482	3	91	down 1,39	cytosol	RNA and DNA binding
Translation elongation factor 1, alpha	7508	nuclear	EEF1A1	gi|31092	3	128	up 2,55	cytosol	GTPase activity
U6 snRNA-associated Sm-like protein LSm8	1205	total lysate	LSM8	gi|7706425	2	85	down 1,65	nucleus	RNA binding
Protein folding and degradation									
Antigen NY-CO-10	3704	nuclear	SDCCAG10	gi|3170184	2	107	down 2,56	nucleus	peptidyl-prolyl cis-trans isomerase activity
Heat shock 70 kDa protein 8, hsc70	2705	nuclear	HSPA8	gi|62897129	14	873	down 2,50	nucleus	ATPase activity
Proteasome activator complex subunit 3	4401	total lysate	PSME3	gi|49456449	7	84	down 3,95	cytosol	protein binding
Protein disulfide isomerase-associated 5	7618	nuclear	PDIA5	gi|5803121	4	130	up 3,03	endoplasmic reticulum	protein disulfide isomerase activity
Protein disulfide-isomerase precursor	1703	total lysate	P4HB	gi|15680282	7	61	up 1,76	endoplasmic reticulum	peptidyl-proline 4-dioxygenase activity
Protein disulfide-isomerase precursor	1704	total lysate	P4HB	gi|15680282	8	235	up 1,62	endoplasmic reticulum	peptidyl-proline 4-dioxygenase activity
T-complex protein 1 subunit delta	8707	total lysate	CCT4	gi|38455427	15	202	down 1,64	cytosol	ATP binding
Ubiquitin	3005	nuclear	RPS27A	gi|229532	3	61	up 2,72	cytosol, nucleus	protein binding
Electron and mitochondrial transport									
Vascular H+ ATPase E1 isoform a	6211	nuclear	ATP6V1E1	gi|4502317	6	266	up 2,25	mitochondrion	ATPase activity
Others (glucide, lipid, nucleotide metabolism, etc…)									
Acetyl-Coenzyme A acetyltransferase 1	8412	nuclear	ACAT1	gi|39795296	2	66	up 2,94	mitochondrion	transferase activity
Aldolase A protein	8401	nuclear	ALDOA	gi|28595	1	75	down 2,27	cytosol	aldehyde-lyase activity
C-1-tetrahydrofolate synthase	6801	nuclear	MTHFD1	gi|115206	4	243	down 5,00	mitochondrion	cyclohydrolase activity
Dihydrolipoamide S-acetyltransferase	2712	nuclear	DLAT	gi|35360	4	169	down 4,17	mitochondrion	transferase activity
Enoyl-CoA hydratase	3208	nuclear	ECH1	gi|70995211	8	612	up 2,38	mitochondrion	isomerase activity
Phosphoserine aminotransferase	7503	total lysate	PSAT1	gi|20141815	4	62	down 2,03	cytosol	transferase activity
Phosphoserine aminotransferase	7504	total lysate	PSAT1	gi|16741698	8	247	down 1,62	cytosol	transferase activity
Miscellaneous									
Autophagy-related protein 16-1	2505	total lysate	ATG16L1	gi|62510482	7	61	down 2,87	autophagic vacuole	protein binding
DAZ associated protein 1	8503	nuclear	DAZAP1	gi|8671754	3	128	up 2,02	nucleus	RNA binding
Androgen-regulated protein 2	2205	total lysate	HN1	gi|7705877	2	53	up 1,80	nucleus	RNA binding
Myosin regulatory light chain 2	107	nuclear	MYL9	gi|20141521	2	151	up 2,22	cytoskeleton	calcium binding

the nuclear proteins, the extraction from 10^9 untreated and RC307-treated cells was obtained with the CelLytic NuCLEAR Extraction Kit (Sigma) following manufacturer's instructions. Then 1% pH 3–10 Ampholyte was added and samples were centrifuged as above to remove the nucleic acids. Both total lysates and nuclear extracts were incubated with 5 mM tributyl phosphine and 20 mM acrylamide for 60 min at room temperature to reduce protein disulphide bonds and alkylate the cysteine thiolic groups. The reaction was blocked by the addition of 10 mM DTT (Sigma) and the samples were collected and stored at −80 °C. Protein concentration was evaluated with DC Protein assay (Bio-Rad, Labs., Hercules, CA, USA) based on the Lowry method.

2.5. Two-dimensional gel electrophoresis

Protein fractionation by 2-DE were performed as previously described [8]. Briefly, 450 µl of each sample (containing 3 mg/ml of protein) was separated by 17 cm pH 3–10 IPG strip, and the total product time×voltage applied was 70000 Vh for each strip. The second dimensional separation was done using 8–18%T gradient SDS-PAGE, applying 40 mA for each gel for 3 min, then 2 mA/gel for 1 h, and 20 mA/gel until the track dye, Bromophenol-Blue, reached the anodic end of the gels. After 2-DE, the proteins were detected by Sypro Ruby.

2.6. Image analysis

The image analysis of the 2D gels replicates was performed by PDQuest software (Bio-Rad), version 7.3. Each gel was analyzed for spot detection, background subtraction and protein spot OD intensity quantification. The gel image showing the higher number of spots and the best protein pattern was chosen as a reference template, and spots in a standard gel were then matched across all gels. Spot quantity values were normalized in each gel dividing the raw quantity of each spot by the total quantity of all the spots included in the standard gel. Two distinct differential analyses were performed, one for total lysates and one for nuclear extracts. In both the experiments gels were divided in two separated groups (control and RC307-treated samples) and, for each protein spot, the average spot quantity value and its variance coefficient in each group were determined. A Student's *t*-test was performed in order to compare the two groups and identify sets of proteins that showed a statistically significant difference with a confidence level of 0.05.

2.7. In-gel digestion

Spots were carefully cut out from 2-D Sypro Ruby stained gels and subjected to in-gel trypsin digestion. Briefly, spots were destained (1×15 min 300 µL wash in 100 mM NH_4HCO_3; 1×15 min 300 µL wash in 50% 100 mM NH_4HCO_3 (v/v), 50% acetonitrile; 1×5 min wash in 100% acetonitrile), and dried at 37 °C. The gel pieces were then swollen in 10 µL of a digestion buffer containing 100 mM NH_4HCO_3 and 20 ng/µL of trypsin (modified porcine trypsin, sequencing grade, Promega, Madison, WI). After 10 min 40 µL of 100 mM NH_4HCO_3 were added to the gel pieces and digestion allowed to proceed at 37 °C. The supernatants were collected and peptides were extracted in an ultrasonic bath for 10 min (twice 50 µL 50% acetonitrile, 50% H_2O with 1% formic acid v/v; once 25 µL of acetonitrile). All the supernatants coming from the different steps were collected in the same tube. Tryptic peptides were dried by vacuum centrifugation and redissolved in 20 µL 0.1% formic acid in water and purified by using ZIP-TIP C18 (Millipore Bedford, MA, USA).

2.8. Peptide sequencing by MALDI-TOF/TOF and nano RP-HPLC-ESI MS/MS

Tryptic peptides for MALDI TOF/TOF analysis were prepared by diluting 1 µl of peptide solution with 1 µl with a saturated solution of α-cyano-4-hydroxy-cinnamic acid (α-CHCA) (10 mg/mL) containing 0.1% TFA and 50% acetonitrile. MS analysis was conducted with a 4800 MALDI TOF/TOF mass spectrometer (Applied Biosystems, Framingham, MA, USA) in positive-ion reflectron mode. For MS analysis, data were acquired automatically after 600 laser shots with a fixed laser intensity of 3200. Each spot sample was calibrated by using an external calibration solution. MS/MS analysis was performed using 1 kV collision energy and air as a collision gas. For peptide sequencing, the 10 most intense precursors were selected. Data were acquired automatically after 1500 laser shots with a fixed laser intensity of 4200. Data were analyzed using GPS Explorer software (Applied Biosystem) and MASCOT software (Matrix Science, London, UK). NCBInr and human were selected as the database and taxonomy, respectively. The remaining peptide mixtures were separated by using a nanoflow-HPLC system (Ultimate; Switchos; Famos; LC Packings, Amsterdam, The Netherlands). A sample volume of 10 µL was loaded by the autosampler onto a homemade 2 cm fused silica pre-column (75 µm I.D.; 375 µm O.D.; Resprosil C18-AQ, 3 µm (Ammerbuch-Entringen, DE) at a flow rate of 2 µL/min. Sequential elution of peptides was accomplished using a flow rate of 200 nL/min and a linear gradient from Solution A (2% acetonitrile; 0.1% formic acid) to 50% of Solution B (98% acetonitrile; 0.1% formic acid) in 40 min over the pre-column in-line with a homemade 10–15 cm resolving column (75 µm I.D.; 375 µm O.D.; Resprosil C18-AQ, 3 µm (Ammerbuch-Entringen, Germany). Peptides were eluted directly into an ion trap Esquire 3000 plus (Bruker-Daltonik, 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. Protein identification was performed by searching in the National Center for Biotechnology Information non-redundant database (NCBInr) using the Mascot program (http://www.matrixscience.com). The following parameters were adopted for database searches: complete propionamide formation on cysteines and partial oxidation of methionines, peptide Mass Tolerance ± 1.2 Da, Fragment Mass Tolerance ± 0.9 Da, missed cleavages 2. For positive identification, the score of the result of [−10×Log(P)] had to be over the significance threshold level ($p<0.05$).

2.9. Protein validation by western blot analysis

For the purpose of this analysis, three biological replicate experiments for the two samples (control and RC307-treated cell line) were analyzed. Protein extracts from HCT116 cell line treated with RC307 for 24 h (2.5 µM) as described above were diluted 1:1 with Laemmli's sample buffer (62.5 mM Tris-HCl, pH 6.8, 25% glycerol, 2% SDS, 0.01% Bromophenol Blue, 5% ß-mercaptoethanol) boiled for 3 min and separated by SDS/polyacrylamide gel electrophoresis (PAGE) on 12% T acrylamide gels in Tris/glycine/SDS buffer. Protein electroblotting and chemiluminescent signal detection were performed as previously described [12]. Briefly, 1-D SDS-PAGE gels were transferred to a PVDF membrane and treated with the respective antibodies at the appropriate dilutions (see the Supplemental Table 1). Bound antibody was detected by enhanced chemiluminescent (ECL) detection kit (Amersham Biosciences Europe) and recorded with X-ray X-Omat AR (Kodak, Rochester, NY, USA) films. Membranes were immunoblotted again with a monoclonal anti-β actin antibody (Sigma-Adrich, 1:4.000) for normalization purposes. The intensity of the chemiluminescence response was measured by scanning films and processing the image using Quantity One software Version 4.4 (Bio-Rad).

2.10. Protein categorization

Gene ontology (GO) lists were downloaded using the tool FatiGO [13] from Babelomics (http://fatigo.bioinfo.cipf.es/), a complete suite of web tools for the functional analysis of groups of genes in high-throughput experiments. Each protein was classified with respect to its cellular component, biological process, and molecular function using GO annotation. When no GO annotation was available, proteins were annotated manually based on literature searches and closely related homologues.

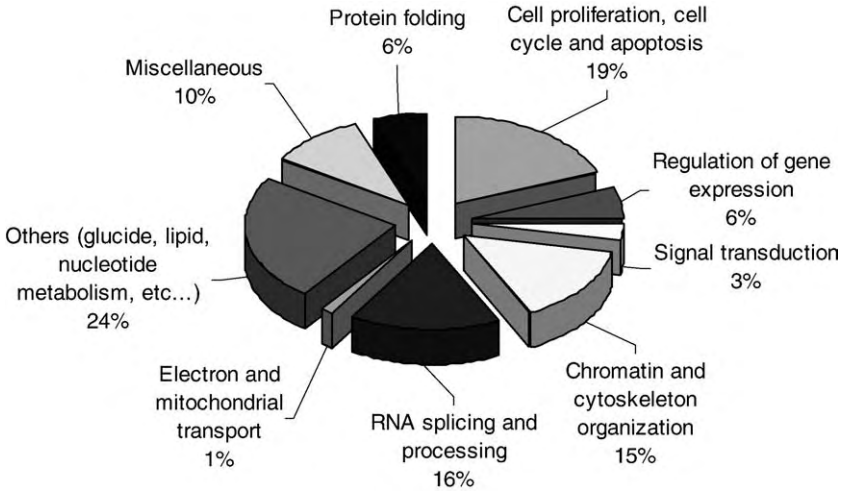

Fig. 3. Proteins GO Categorization. Distribution of the identified proteins according to the biological process in which they are involved. Assignments were made on the basis of information provided by Gene ontology (GO) lists downloaded using the tool FatiGO [13] from Babelomics (http://fatigo.bioinfo.cipf.es/).

3. Results

3.1. Effects of RC307 on HCT116 cells

Sensitivity of HCT116 cells to the HDAC inhibitor RC307 was assessed by growth-inhibition assay following exposure to the drug for 72 h. The compound exhibited a marked capability to inhibit proliferation, the IC_{50} value being in the micromolar range, i.e., 1.14±0.15 μM (n=6) (Fig. 1A). Such an effect was associated with induction of apoptosis as shown by TUNEL assay in cells exposed for 72 h to a concentration of the compound producing an inhibition of cell growth of around 80% (Fig. 1B). Indeed, an appreciable amount of cells with DNA fragmentation was documented by labelling of DNA strand breaks. The occurrence of apoptotic cell death was also supported by cytofluorimetric analysis of cell cycle showing the appearance of a sub-G1 peak after a 24-h drug exposure. Under such conditions, the antiproliferative treatment produced a marked accumulation in the G2 phase of the cell cycle (Fig. 1C).

3.2. HCT116 Cell Line 2-DE protein pattern analysis

In order to examine the effect of HDAC inhibition by RC307 on protein expression in a human colorectal carcinoma cell line (HCT116), five replicated maps for each experimental group (controls and RC307-treated) were performed. The separated protein spots were visualized on 2D gels by Sypro Ruby staining, which allows good reproducibility and protein spot quantification for comparison analysis. The 2D gel of total lysates showed a total of 480±21 and 474±22 protein spots, for the control and RC307-treated cells, respectively; while the 2D gels of nuclear extracts showed a total of 774±19 and 781±15 protein spots, for the control and RC307-treated cells, respectively. By PDQuest analysis of the 2D gel replicates, we measured differential protein expression between control and RC307-treated cells. A total of 48 and 46 different spots were found to be modulated by RC307 in total lysates and nuclear extracts, respectively. In particular, 27 spots were found to be up-regulated and 21 down-regulated in the total lysates; while 24 up-regulated and 22 down-regulated in the nuclear extracts. Fig. 2A and B show representative 2D gels of total lysate and nuclear proteome of HCT116 cell line, respectively, together with some of the corresponding differentially expressed spots ($p<0.05$) between control and RC307-treated samples. Spots selected as regulated from the differential analysis were subjected to MALDI-TOF/TOF and RP-HPLC-ESI-MS/MS analysis for protein identification. The unique differentially expressed proteins identified were 66. In Table 1, the identity of the successfully identified proteins corresponding to up- or down-regulated spots (confidence level of 0.05) are shown, together with the standard spot number (SSP), the MS identification parameters and the indication of their gene ontology (GO) annotation (cellular component, biological process and molecular function). Not all the differentially expressed spots were identified because of their relative low concentrations. Fig. 3 shows as a pie chart the distribution of the identified proteins catalogued according to the biological process in which they are involved. To validate the findings obtained by 2-DE, the regulation in level of expression of four candidate proteins (cytokeratin 8, cytokeratin 18, alpha enolase and stathmin) were further investigated by immunoblot analysis. Western blot results are shown in Fig. 4 including the relative protein expression in both samples normalized to β-actin signal intensity as an internal loading control. Trends of changes in the same direction as those detected in the 2D gel analyses were detected for all the four proteins. The quantitative difference between the results obtained by 2D electrophoresis and by Western blot suggested that most changes detected by the former technique specifically involve post-translationally modified forms, which can be only separated in 2D maps (Fig. 2A and B).

4. Discussion

In this study, we report that RC307, a histone deacetylase inhibitor, determined cell growth inhibition, G2 accumulation and a moderate induction of apoptosis of HCT116 colon cancer cell line. To analyze the molecular mechanisms involved in cell growth inhibition and the potential target genes for aberrant acetylation, we performed a proteomic analysis of HCT116 after treatment with RC307. The differentially expressed proteins identified by 2-D proteome analysis are related to various cellular programs involving, e.g., proliferation, cell cycle and apoptosis regulation, gene expression, as well as chromatin and cytoskeleton organization (Table 1). Interestingly, some of these identified proteins (such as Glutathione S-transferase pi, Peroxiredoxin 1, TNF receptor-associated protein 1, Profilin I, Annexin II, Cytokeratin 8 and 18, Lamin A/C, and U6 snRNA-associated Sm-like protein LSm8) correspond to those that we reported as modulated by trichostatin A (showing the same trend of variation) in our previous studies [6–8].

The cellular effect of RC307 could be explained by modulation of expression of proteins that have been implicated in cell proliferation, cell cycle and apoptosis. In the present study, we showed that RC307 down-regulates Galectin 3 (LGALS3). LGALS3 is a beta-galactoside-binding protein whose expression has been correlated with progression and metastasis in colon cancer and has been reported to represent a potent prognostic marker in colorectal cancer [14,15]. Furthermore, the interaction between LGALS3 and Bcl-2 suggests that LGASL3 is involved in the inhibition of apoptosis [16,17]. It is possible that RC307-mediated LGALS3 reduction could help to induce the apoptotic response of HCT116 cells. We also found that RC307 modulates the Glutathione S-transferase pi (GSTP1), which belongs

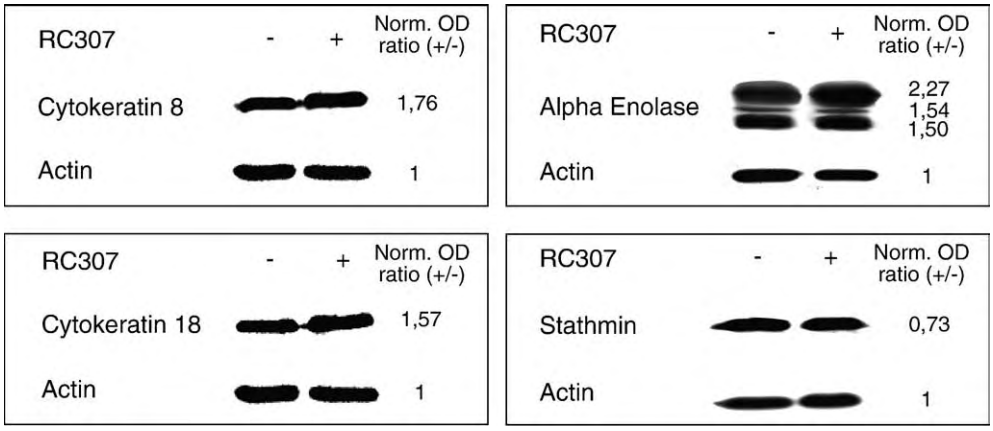

Fig. 4. Protein expression validation by Western blot analysis. Western blot analyses for four proteins showing regulation in RC307 treated HCT116 cell line by proteomic analysis. Film images with the relative protein expression normalized to actin signal intensity as an internal control. Western blot images were captured by GS710 densitometer (Bio-Rad) and analyzed by QuantityOne software to calculate the band intensities (OD).

to a family of isoenzymes known to inactivate damaging electrophilic compounds by catalyzing their conjugation to reduced glutathione. In most experimental systems, over-expression of GSTP1 is associated with increased resistance to anticancer agents [18]. Although there are contrasting reports regarding the association of GSTP1 upregulation with the apoptotic response induced in colon cancer cells by the HDAC inhibitor butyrate [19,20], our group has previously shown that the apoptotic response induced in pancreatic cancer cell lines by TSA is related to GSTP1 overexpression [7,8]. These findings are intriguing since the lack of GSTP1 expression is associated with methylation in the gene promoter [21], and methylation-induced gene silencing has been associated with histone deacetylation. Our findings showing upregulation of GSTP1 by RC307 are in keeping with such reports, as inhibition of HDAC could relieve gene silencing. Another protein modulated by RC307 was Peroxiredoxin 1 (PRDX1), a ubiquitously expressed member of a family of antioxidant proteins induced by reactive oxygen species. PRDX1 has been reported to inhibit the c-myc oncogene and to act as a tumour suppressor [22,23]. Moreover, other HDAC inhibitors (i.e., FK288 and TSA) have been shown to up-regulate PRDX1 in esophageal and pancreatic cancer cell lines in which apoptosis has been documented [24,6–8]. Consistent with this, here we show that also RC307 enhances the level of the tumour suppressor PRDX1. Another protein that we found up-regulated by RC307 was the Heat shock 90 kDa protein 1 beta (HSP90AB1), an ATP-dependent chaperone that plays a central role in regulating the stabilization, activation, and degradation of a range of proteins including the products of oncogenes. HSP90AB1 has become an attractive target for novel cancer therapeutic agents since its inhibition disrupts multiple cancer-causing pathways simultaneously [25,26]. Recently, it has been reported HCT116 cell growth inhibition after treatment with different hsp90 inhibitors [27]. Since we found that RC307 reduces the level of HSP90AB1 expression it is conceivable that the event is a consequence of HSP90ABI acetylation. Another protein regulated by RC307 was Stathmin (STMN1), a microtubule destabilizing protein previously described as being negatively regulated by p53 [28] and highly expressed in several tumours. Recently, it has been demonstrated that downregulation of STMN1 by siRNA in osteosarcoma cell lines induces G(2)/M cell cycle arrest and apoptotic cell death [29]. Accordingly, the reduction of STMN1 level increases the responsiveness of tumour cells to treatment with chemotherapeutic agents [30]. Consistent with this, here we showed that RC307 down-regulates STMN1, and we were able to validate this finding by immunoblot analysis. Collectively, all these findings suggest that cell response of HCT116 cells to RC307 involves proteins that play a key role in colon cancer cell proliferation, cell cycle arrest and apoptosis.

The cellular effect of RC307 could also be explained by the modulation of proteins involved in gene expression regulation. We identified Alpha-enolase (ENO1) as a protein regulated by RC307. ENO1 is a bifunctional gene encoding both a glycolytic enzyme (ENO1, 48 kDa), and a transcription factor (MBP-1, 37 kDa) which binds and represses c-myc gene playing an important role in cancer cell growth inhibition [31]. Surprisingly, it has been demonstrated that also the longer form (ENO1) alone has equal or stronger effect on the apoptosis induction, as MBP-1 does [32]. Consistent with this, we found that RC307 upregulates ENO1, a finding that was validated by immunoblot analysis. Here we report that RC307 modulates also cytokine-inducible 29-kDa protein (CIP29). CIP29 is a cytokine regulated nuclear protein that binds both double- and single-stranded DNA. This nuclear protein is involved in normal and cancer cell proliferation [33]. Interestingly, it has been reported that CIP29 has a growth inhibitory effect associated with induction of apoptosis [34]. Accordingly, we found that RC307 upregulates CIP29 in colon cancer cell lines. It is possible that RC307-mediated ENO1 and CIP29 upregulation could enhance the apoptotic response of HCT116 cells.

The mechanism of action of RC307 also involves the regulation of proteins related to chromatin and cytoskeleton organization. In particular, among the cytoskeleton related proteins, we found cytokeratins isoforms 8 (KRT8) and 18 (KRT18) as upregulated by RC307. KRT8 and KRT18 are the major components of intermediary filaments of simple or single layer epithelia, such as those of the intestine. While both isoforms are essential for maintaining structural integrity, there is accumulating evidence indicating that they also exert non-mechanical functions. Indeed, the expression of KRT8 and KRT18 has been related to poor clinical prognosis [35]. Moreover, KRT8 and KRT18 upregulation has been associated with the sensitization to apoptosis, induced by cisplatin [36], roscovitine [37], and Fas and TNF [38]. It has also been reported that KRT8 and KRT18 form intermediate filaments that are required by hsp90β (HSP90AB1) for its chaperone activity [39]. Surprisingly, KRT8 and KRT18 have been shown to be upregulated in HCT116 cell line after inhibition of the HSP90AB1 [40]. Based on such reports, it is conceivable that increases in KRT8 and KRT18 level of expression that we reported, are related to the RC307-mediated HSP90AB1 downregulation, and may contribute to HCT116 cellular growth inhibition induced by the HDAC inhibitor RC307.

5. Conclusion

In conclusion, our results indicating modulation of proteins belonging to different pathways (i.e., in proliferation, cell cycle and apoptosis regulation, gene expression, as well as chromatin and cytoskeleton organization) by a novel HDAC inhibitor support the interest of proteomic approaches in defining the mechanism of action of novel anticancer drugs. Among the most interesting RC307-regulated proteins we identified Galectin 3, Peroxiredoxin 1, Heat shock 90 kDa protein 1 beta, Alpha-enolase and cytokeratins 8 and 18. The specific role of each of these proteins in cell response to RC307 remains to be defined.

Acknowledgments

This work was supported in part by grants from MIUR (FIRB) and the Human Frontier Science Program (to AV); by CIB (Consorzio Interuniversitario per le Biotecnologie) and MIUR (PRIN 2007) (to LZ); by AIRC (to FZ) and by Fondazione Cariplo (to FZ and PGR).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbapap.2008.04.022.

References

[1] A.J.M. De Ruijter, A.H. van Gennip, H.N. Caron, S. Kemp, B.P. van Kuilenburg, Histone deacetylases (HDACs): characterization of the classical HDAC family, Biochem. J. 370 (2003) 737–749.
[2] R.W. Johnstone, J.D. Licht, Histone deacetylase inhibitors in cancer therapy: is transcription the primary target? Cancer Cell 4 (2003) 13–18.
[3] R. Brown, G. Strathdee, Epigenomics and epigenetic therapy of cancer, Trends Mol. Med. 8 (2002) 43–48.
[4] R.W. Johnstone, Histone-deacetylase inhibitors: novel drugs for the treatment of cancer, Nat. Rev. Drug Discov. 12 (2002) 294–300.
[5] F. McLaughlin, N.B. La Thangue, Histone deacetylase inhibitors open new doors in cancer therapy, Biochem. Pharmacol. 68 (2004) 1139–1144.
[6] D. Cecconi, A. Scarpa, M. Donadelli, M. Palmieri, M. Hamdan, H. Astner, P.G. Righetti, Proteomic profiling of pancreatic ductal carcinoma cell lines treated with trichostatin-A, Electrophoresis 24 (2003) 1871–1878.
[7] D. Cecconi, M. Donadelli, A. Scarpa, A. Milli, M. Palmieri, M. Hamdan, J. Rappsilber, L.B. Areces, P.G. Righetti, Proteomic analysis of pancreatic ductal carcinoma cells after combined treatment with gemcitabine and trichostatin A, J. Proteome Res. 4 (2005) 1900–1916.
[8] D. Cecconi, M. Donadelli, S. Rinalducci, L. Zolla, A. Scarpa, M. Palmieri, P.G. Righetti, Proteomic analysis of pancreatic endocrine tumour cell lines treated with the histone deacetylase inhibitor trichostatin A, Proteomics 7 (2007) 1644–1653.
[9] J-H. Park, Y. Jung, T.Y. Kim, S.G. Kim, H-S. Jong, J.W. Lee, et al., Class I histone deacetylase-selective novel sunthetic inhibitors potently inhibit human tumor proliferation, Clin. Cancer Res. 10 (2004) 5271–5281.
[10] T. Beckers, C. Burkhardt, H. Wieland, P. Gimmnich, T. Ciossek, T. Maier, K. Sanders, Distinct pharmacological properties of second generation HDAC inhibitors with the benzamide or hydroxamate head group, Int. J. Cancer 121 (2007) 1138–1148.

[11] P. Atadja, L. Gao, P. Kwon, N. Trogani, H. Walker, et al., Selective growth inhibition of tumor cells by a novel histone deacetylase inhibitor, NVP-LAQ824, Cancer Res. 64 (2004) 689–695.
[12] D. Cecconi, S. Mion, H. Astner, E. Domenici, P.G. Righetti, L. Carboni, Proteomic analysis of rat cortical neurons after fluoxetine treatment, Brain Res. 1135 (2007) 41–51.
[13] F. Al-Shahrour, P. Minguez, J. Tárraga, D. Montaner, E. Alloza, J.M. Vaquerizas, et al., BABELOMICS: a systems biology perspective in the functional annotation of genome-scale experiments, Nucleic Acids Res. 34 (2006) W472–W476.
[14] K. Endo, S. Kohnoe, E. Tsujita, A. Watanabe, H. Nakashima, H. Baba, Y. Maehara, Galectin-3 expression is a potent prognostic marker in colorectal cancer, Anticancer Res. 25 (2005) 3117–3121.
[15] C. Greco, R. Vona, M. Cosimelli, P. Matarrese, E. Straface, P. Scordati, D. Giannarelli, V. Casale, D. Assisi, M. Mottolese, A. Moles, W. Malorni, Cell surface overexpression of galectin-3 and the presence of its ligand 90 k in the blood plasma as determinants in colon neoplastic lesions, Glycobiology 14 (2004) 783–792.
[16] S. Akahani, P. Nangia-Makker, H. Inohara, H.R. Kim, A. Raz, Galectin-3: a novel antiapoptotic molecule with a functional BH1 (NWGR) domain of Bcl-2 family, Cancer Res. 57 (1997) 5272–5276.
[17] R.Y. Yang, D.K. Hsu, F.T. Liu, Expression of galectin-3 modulates T-cell growth and apoptosis, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 6737–6742.
[18] D.M. Townsend, K.D. Tew, The role of glutathione-Stransferase in anti-cancer drug resistance, Oncogene 22 (2003) 7369–7375.
[19] M.N. Ebert, G. Beyer-Sehlmeyer, U.M. Liegibel, T. Kautenburger, T.W. Becker, B.L. Pool-Zobel, Butyrate induces glutathione S-transferase in human colon cells and protects from genetic damage by 4-hydroxy-2-nonenal, Nutr. Cancer 41 (2001) 156–164.
[20] B.L. Pool-Zobel, V. Selvaraju, J. Sauer, et al., Butyrate may enhance toxicological defence in primary, adenoma and tumor human colon cells by favourably modulating expression of glutathione S-transferases genes, an approach in nutrigenomics, Carcinogenesis 6 (2005) 1064–1076.
[21] W.H. Lee, R.A. Morton, J.I. Epstein, J.D. Brooks, P.A. Campbell, G.S. Bova, W.S. Hsieh, W.B. Isaacs, W.G. Nelson, Cytidine methylation of regulatory sequences near the pi-class glutathione S-transferase gene accompanies human prostatic carcinogenesis, Proc. Natl. Acad. Sci. U. S. A. 91 (1994) 11733–11737.
[22] C.A. Neumann, Q. Fang, Are peroxiredoxins tumor suppressors? Curr. Opin. Pharmacol. 7 (2007) 375–380.
[23] R.A. Egler, E. Fernandes, K. Rothermund, S. Sereika, N. de Souza-Pinto, P. Jaruga, M. Dizdaroglu, E.V. Prochownik, Regulation of reactive oxygen species, DNA damage, and c-Myc function by peroxiredoxin 1, Oncogene 24 (2005) 8038–8050.
[24] I. Hoshino, H. Matsubara, N. Hanari, M. Mori, T. Nishimori, Y. Yoneyama, Y. Akutsu, H. Sakata, K. Matsushita, N. Seki, T. Ochiai, Histone deacetylase inhibitor FK228 activates tumor suppressor Prdx1 with apoptosis induction in esophageal cancer cells, Clin. Cancer Res. 11 (2005) 7945–7952.
[25] S.B. Cullinan, L. Whitesell, Heat shock protein 90: a unique chemotherapeutic target, Semin Oncol. 33 (2006) 457–465.
[26] C. Dai, L. Whitesell, HSP90: a rising star on the horizon of anticancer targets, Future Oncol. 1 (2005) 529–540.
[27] S.Y. Sharp, C. Prodromou, K. Boxall, M.V. Powers, J.L. Holmes, G. Box, T.P. Matthews, K.M. Cheung, A. Kalusa, K. James, A. Hayes, A. Hardcastle, B. Dymock, P.A. Brough, X. Barril, J.E. Cansfield, L. Wright, A. Surgenor, N. Foloppe, R.E. Hubbard, W. Aherne, L. Pearl, K. Jones, E. McDonald, F. Raynaud, S. Eccles, M. Drysdale, P. Workman, Inhibition of the heat shock protein 90 molecular chaperone in vitro and in vivo by novel, synthetic, potent resorcinylic pyrazole/isoxazole amide analogues, Mol. Cancer Ther. 6 (2007) 1198–1211.
[28] J.I. Johnsen, O.N. Aurelio, Z. Kwaja, G.E. Jörgensen, N.S. Pellegata, R. Plattner, E.J. Stanbridge, J.F. Cajot, p53-mediated negative regulation of stathmin/Op18 expression is associated with G(2)/M cell-cycle arrest, Int. J. Cancer. 88 (2000) 685–691.
[29] R. Wang, K. Dong, F. Lin, X. Wang, P. Gao, S.H. Wei, S.Y. Cheng, H.Z. Zhang, Inhibiting Proliferation and Enhancing Chemosensitivity to Taxanes in Osteosarcoma Cells by RNA Interference-Mediated Downregulation of Stathmin Expression, Mol. Med. 13 (2007) 567–575.
[30] S. Singer, V. Ehemann, A. Brauckhoff, M. Keith, S. Vreden, P. Schirmacher, K. Breuhahn, Protumorigenic overexpression of stathmin/Op18 by gain-of-function mutation in p53 in human hepatocarcinogenesis, Hepatology 46 (2007) 759–768.
[31] K.C. Sedoris, S.D. Thomas, D.M. Miller, c-myc promoter binding protein regulates the cellular response to an altered glucose concentration, Biochemistry 46 (2007) 8659–8668.
[32] K. Ejeskär, C. Krona, H. Carén, F. Zaibak, L. Li, T. Martinsson, P.A. Ioannou, Introduction of in vitro transcribed ENO1 mRNA into neuroblastoma cells induces cell death, BMC Cancer 5 (2005) 161.
[33] S. Fukuda, D.W. Wu, K. Stark, L.M. Pelus, Cloning and characterization of a proliferation-associated cytokine-inducible protein, CIP29, Biochem. Biophys. Res. Commun. 292 (2002) 593–600.
[34] S. Fukuda, L.M. Pelus, Growth inhibitory effect of Hcc-1/CIP29 is associated with induction of apoptosis, not just with G2/M arrest, Cell. Mol. Life Sci. 62 (2005) 1526–1527.
[35] T. Fillies, R. Werkmeister, J. Packeisen, B. Brandt, P. Morin, D. Weingart, U. Joos, H. Buerger, Cytokeratin 8/18 expression indicates a poor prognosis in squamous cell carcinomas of the oral cavity, BMC Cancer. 6 (2006) 10.
[36] H.K. Parekh, H. Simpkins, The differential expression of cytokeratin 18 in cisplatin-sensitive and -resistant human ovarian adenocarcinoma cells and its association with drug sensitivity, Cancer Res. 55 (1995) 5203–5206.
[37] B. Schutte, M. Henfling, F.C. Ramaekers, DEDD association with cytokeratin filaments correlates with sensitivity to apoptosis, Apoptosis 11 (2006) 1561–1572.
[38] S. Gilbert, A. Loranger, N. Marceau, Keratins modulate c-Flip/extracellular signal-regulated kinase 1 and 2 antiapoptotic signaling in simple epithelial cells, Mol. Cell. Biol. 24 (2004) 7072–7081.
[39] M.D. Galigniana, J.L. Scruggs, J. Herrington, M.J. Welsh, C. Carter-Su, P.R. Housley, W.B. Pratt, Mol, Endocrinol. 12 (1998) 1903–1913.
[40] P.A. Clarke, I. Hostein, U. Banerji, F.D. Stefano, A. Maloney, M. Walton, I. Judson, P. Workman, Gene expression profiling of human colon cancer cells following inhibition of signal transduction by 17-allylamino-17-demethoxygeldanamycin, an inhibitor of the hsp90 molecular chaperone, Oncogene 19 (2000) 4125–4133.

RESEARCH ARTICLE

Signal transduction pathways of mantle cell lymphoma: A phosphoproteome-based study

Daniela Cecconi[1], Alberto Zamò[2]*, Elena Bianchi[1], Alice Parisi[2], Stefano Barbi[2], Alberto Milli[1], Sara Rinalducci[3], Andreas Rosenwald[4], Elena Hartmann[4], Lello Zolla[3] and Marco Chilosi[2]*

[1] Laboratorio di Proteomica, Dipartimento Scientifico e Tecnologico, University of Verona, Verona, Italy
[2] Dipartimento di Patologia, Sezione di Anatomia Patologica, University of Verona, Verona, Italy
[3] Dipartimento di Scienze Ambientali, University of Tuscia, Viterbo, Italy
[4] Department of Pathology, University of Wuerzburg, Wuerzburg, Germany

Mantle cell lymphoma (MCL) is an incurable hematologic malignancy whose pathogenesis is only partly understood. The aim of the present study was to define a "core phosphoproteome" in MCL cell lines that is representative of primary MCL in order to improve knowledge of the signal transduction pathways involved in its tumorigenesis. We have analyzed phosphorylated proteins in several MCL cell lines by immobilized metal affinity chromatography and separation by 2-D PAGE, followed by RP-HPLC coupled with MS/MS identification. These data were correlated with information on copy number gains obtained by SNP-chip analysis. Several of the proteins identified could be linked to a specific signal transduction pathway, and have been recently recognized as important players in MCL pathogenesis, such as nuclear factor-kappaB (NF-κB) and phosphoinositide-3 kinase-mammalian target of rapamycin (PI3K-mTOR). However, our data also implicate a number of novel proteins and pathways in the pathobiology of MCL, one of which is mitochondrial signaling. A second-level analysis identified MAPK1, CK2, CK1, PKCzeta, and PKCepsilon as candidate upstream molecules. Our study provides new insights in MCL pathogenesis and helps to form the basis for testing new target-specific therapeutics.

Received: January 28, 2008
Revised: May 6, 2008
Accepted: June 3, 2008

Keywords:
Mantle cell lymphoma / Phosphoproteins / Signal transduction

1 Introduction

Mantle cell lymphoma (MCL) constitutes about 5% of non-Hodgkin lymphomas (NHL) [1]. It is considered a B-cell lymphoma originating from mature but naive B-cells. It is characterized by a specific t(11;14) translocation, juxtaposing the CCND1 locus on chromosome 11 to the Ig enhancer on chromosome 14, thus resulting in hyperexpression of cyclin D1. This event appears to be necessary but not sufficient for neoplastic transformation [2], and other abnormalities are involved in the pathogenesis of MCL. Since current therapies for MCL are not satisfactory, and relapse is frequent, newer treatment modalities are needed. In this regard, significant efforts have been devoted to the use of modern high-throughput techniques, such as gene-expression profiling [3]. Nonetheless, expression-based techniques have the disadvantage of being one step away from the targets of most drugs, namely proteins. Previous studies have carried proteomic analysis of MCL tissues or cell lines [4, 5].

Correspondence: Dr. Alberto Zamò, Dipartimento di Patologia, Sezione di Anatomia Patologica, Verona, Strada le Grazie 8, 37134, Verona, Italy
E-mail: alberto.zamo@univr.it
Fax: +39-045-802-7136

Abbreviations: eEF2, eukaryotic elongation factor 2; **eIF3,** eukaryotic translation initiation factor 3; **HSPA5,** heat shock protein 5; **IMAC,** immobilized metal affinity chromatography; **GO,** gene ontology; **MCL,** mantle cell lymphoma; **mTOR,** mammalian target of rapamycin; **NF-κB,** nuclear factor-kappaB; **PI3K,** phosphoinositide-3 kinase

* These authors contributed equally and share first authorship.

In the present investigation, we attempted to reduce sample complexity by focusing only on putatively active, phosphorylated proteins. Protein phosphorylation of serine, threonine, and tyrosine residues is a fundamental PTM that regulates cell signaling, oncogenesis, apoptosis, and immune disorders. Phosphoproteins are labile and often present in low quantities, and thus their investigation requires specific prefractionation strategies. The phosphoproteome has been conventionally explored by the incorporation of radiolabeled phosphate into proteins followed by 2-DE. More recently other experimental strategies have become available such as: MS [6, 7], specific capture of phosphopeptides by antibodies, β-elimination of phosphate from serine or threonine followed by attack of an affinity tag [8–13], direct staining of phosphorylated proteins in polyacrylamide gels, immunoprecipitation, and immobilized metal affinity chromatography (IMAC) [14, 15]. Herein we used a commercially available IMAC kit (PhosphoProtein Purification Kit, Qiagen), to obtain complete and reproducible separation of the phosphorylated protein fraction with the aim of identifying a "core phosphoproteome" that might be cell line independent, and therefore representative of primary MCL.

For this purpose, we analyzed the global phosphoproteome profiles of seven different MCL cell lines (Granta-519, Jeko-1, MAVER-1, NCEB-1, Rec-1, UPN-1, and UPN-2), and a subset of highly expressed phosphoproteins was isolated. Phosphoproteins were characterized using bioinformatics, which allowed the assignment of number and position of potential phosphorylation sites, the search for kinases and phosphorylation-dependent binding motifs, and the classification of the phosphoproteins with respect to their Gene Ontology (GO) annotations. Some of the proteins belong to signal transduction pathways that are relevant to tumorigenesis, such as nuclear factor-kappaB (NF-κB), mammalian target of rapamycin (mTOR), and mitochondrial signaling. On the same cell lines, we also performed high-density SNP-chip analysis, which was correlated with the phosphoproteins identified. Several proteins were found that showed a correspondent copy number gain at the DNA level. The identification of these proteins allowed us to formulate rational hypotheses about their role as candidate "drivers" of MCL pathogenesis.

2 Materials and methods

2.1 Cell culture

The MAVER-1 cell line was established in our laboratories and previously described [16]. Cell lines Granta-519 [17], Rec-1 [18], and Jeko-1 [19] were obtained from DSMZ (Braunschweig, Germany). UPN-1 [20], UPN-2 [20], and NCEB-1 [21] were provided by Professor Elias Campo (Hospital Clinic, University of Barcelona, Spain). Cell lines were cultured in RPMI 1640 medium containing 10% FCS (BioWhittaker, Walersville, MD, USA), and gathered at plateau phase ($1.5–2.0 \times 10^6$ cells/mL) by centrifugation. Cells were washed twice in sterile 0.9% NaCl followed by freezing in liquid nitrogen.

2.2 IMAC

A phosphoprotein purification kit (Qiagen, Valencia, CA, USA) was utilized for purifying phosphorylated proteins. Cells (1×10^7) were harvested and processed following the manufacturer's instructions. Briefly, cells were gently lysed in a buffer containing 0.25% w/v CHAPS, Benzonase (a DNase/RNase), and a mixture of protease inhibitors to prevent proteolytic degradation. Subsequently, lysates were added to the ready-to-use columns and the flow-through fractions (containing nonphosphorylated proteins) were collected. After a wash step, proteins carrying phosphate groups were eluted in PBS. Free phosphate in the elution buffer inhibits phosphatase activity, and therefore stabilizes the phosphorylation status of the eluted fractions during downstream processing and storage.

2.3 High-resolution 2-DE

Phosphorylated protein fractions were concentrated and desalted before 2-DE by ultrafiltration (Nanosep® Ultrafiltration columns 10 kDa cut-off) and addition of 450 μL 10 mM Tris-Cl, pH 7.0. Protein fractions were then precipitated by acetone/methanol (8:1 v/v) and the pellets resuspended in 2-D solubilizing solution: 7 M urea (Sigma–Aldrich, St. Louis, MO, USA), 2 M thiourea (Sigma), 3% CHAPS (Sigma), 20 mM Tris (Sigma), 1% pH 3–10 Ampholyte (Fluka, Buchs, SG, Switzerland), and protease inhibitor cocktail tablet (Complete Mini; Roche, Basel, Switzerland). After 1 h samples were centrifuged for 10 min at $10\,000 \times g$ at 4°C to remove nucleic acids complexed with ampholytes. Supernatants were incubated with 5 mM tributylphosphine and 20 mM acrylamide for 60 min at room temperature to reduce protein disulfide bonds and alkylate sulfhydryl groups. The reaction was blocked by the addition of 10 mM DTT and protein samples were collected and stored at −80°C. Protein concentration was evaluated with the DC Protein assay (BioRad, Hercules, CA, USA). Seventeen centimeter long, pH 3–10 IPG strips (BioRad) were rehydrated for 8 h with 450 μL 2-D solubilizing solution (7 M urea, 2 M thiourea, 3% CHAPS, and 20 mM Tris) containing 2.5 mg/mL total phosphoprotein. IEF was performed using a Multiphor II apparatus (GE Healthcare BioSciences, Little Chalfont, UK). The total product time × voltage applied was 70 000 V·h for each strip at 20°C. For the second dimension, IPG strips were equilibrated for 26 min in a solution of 6 M urea, 2% SDS, 20% glycerol, 375 mM Tris-HCl, pH 8.8. The IPG strips were then laid on an 8–18%T gradient SDS-PAGE with 0.8% agarose in Tris/glycine/SDS running buffer (192 mM glycine, 0.1% SDS and Tris to pH 8.3). The second dimension was run in a Protean Plus Dodeca cell (BioRad)

with Tris/glycine/SDS running buffer. The electrophoresis was conducted with continuous cooling and mixing by setting a current of 40 mA for each gel for 3 min, then 2 mA/gel for 1 h, and 20 mA/gel until the track dye, bromophenol blue, reached the anodic end of the gels. Protein spots were revealed with Sypro Ruby stain. Gels were incubated in a fixing solution containing 40% ethanol and 10% acetic acid for 30 min followed by overnight staining in a ready-to-use Sypro Ruby solution. Destaining was performed in 10% methanol and 7% acetic acid for 1 h, followed by rinsing at least 3 h in distilled water. Detection of phosphoproteins by Pro-Q Diamond staining (Molecular Probes) after separation on 2-DE gels was performed following the manufacturer's instructions. In brief, 2-DE gels were fixed in solution containing 50% methanol and 10% acetic acid, washed with several changes of water to remove SDS, and stained with the Pro-Q Diamond dye. After destaining, gel images were recorded using a Typhoon 9600 fluorescence scanner. The spots consistently stained with Pro-Q Diamond dye were considered as putative phosphoproteins.

2.4 Protein pattern differential analysis

Gels were scanned using a BioRad VersaDoc 1000 imaging system. 2-D gel analysis was performed by PDQuest software version 7.3 (BioRad). Each gel was analyzed for spot detection, background subtraction, and protein spot OD intensity quantification (spot quantity definition). The gel image showing the highest number of spots and the best protein pattern was chosen as a reference template, and spots in a standard gel were then matched across all gels. Spot quantity values were normalized in each gel dividing the raw quantity of each spot by the total quantity of all the spots included in the standard gel. For each protein spot, the average spot quantity value and its variance coefficient in each group were determined. The average spot intensities of each cell line were ranked in decreasing order, and the 100 spots with the highest average spot intensity value were chosen for the identification.

2.5 In-Gel digestion

Spots were selectively excised from the gel using the Proteome Works robotic Spot cutter (BioRad), which allows for reproducible processing of gels. Excised spots were then subjected to in-gel trypsin digestion. Spots were destained (15 min wash in 300 μL 100 mM NH_4HCO_3; 15 min wash in 300 μL 50% 100 mM NH_4HCO_3 v/v, 50% ACN; 5 min wash in 100% ACN) and dried at 37°C. The gel pieces were swollen in 10 μL of digestion buffer containing 100 mM NH_4HCO_3 and 40 ng/μL trypsin (Promega, Madison, WI, USA). After 10 min, 40 μL of 100 mM NH_4HCO_3 was added to the gel pieces and digestion was performed at 37°C overnight. The supernatants were collected, peptides were extracted in an ultrasonic bath for 10 min (twice in 50 μL 50% ACN, 50% H_2O with 1% formic acid v/v; once in 25 μL ACN), and the supernatants were pooled. Tryptic peptides were vacuum-dried with centrifugation, resuspended in 20 μL 0.1% formic acid and purified using a ZIP-TIP C18 column (Millipore, Bedford, MA, USA).

2.6 Peptide sequencing by nano-RP-HPLC-ESI-MS/MS

Peptide mixtures were analyzed by nanoflow HPLC: a 10 μL sample was loaded onto a homemade 2 cm fused-silica precolumn (75 μm id, 375 μm od, packed with Reprosil C18-AQ, 3 μ/m, at a flow rate of 2 μL/min. Sequential elution of peptides was accomplished using a flow rate of 200 nL/min and a linear gradient from solution A (2% ACN, 0.1% formic acid) to 50% of solution B (98% ACN, 0.1% formic acid) over 40 min. Desorbed peptides were injected directly into a Q-TOF microhybrid mass spectrometer (model microTOF-Q, Bruker-Daltonik) equipped with a modified ESI ion source (the spray capillary was a fused-silica capillary, 0.090 mm od, 0.020 mm id). The mass spectrometer was operated in the positive ion MS mode, and data-dependent analysis was employed for survey scans (m/z 350–1500) to choose up to three most intense precursor ions. For CID mass spectrometric (MS/MS) analysis, collision energies were chosen automatically as a function of m/z and charge. The collision gas was argon. The temperature of the heated sample source was 180°C, and the electrospray voltage was 4000 V. External mass calibration in quadratic regression mode using sodium formiate resulted in mass errors of typically 5 ppm in the m/z range 50–2000. Data searching was accomplished with MASCOT software (Matrix Sciences, London, UK) using the following constraints: only tryptic peptides up to two missed cleavage sites were allowed; ±0.25 Da tolerance for both MS and MS/MS fragment ions; complete propionamide formation on cysteines and partial oxidation of methionines were specified.

2.7 Protein phosphorylation status validation by Western blot analysis

Protein electroblotting and chemiluminescent signal detection were performed as previously described [22]. Briefly, 1-D SDS-PAGE gels were transferred to a PVDF membrane and treated with the respective antiphosphoprotein antibodies at the appropriate dilutions. Bound antibody was detected with the enhanced chemiluminescent (ECL) detection kit (Amersham Biosciences Europe) and recorded on X-Omat AR film (Kodak, Rochester, NY, USA). Membranes were immunoblotted again with a monoclonal anti-β actin antibody (Sigma–Adrich, 1:4.000) for normalization.

2.8 Classification of identified phosphoproteins and phosphorylation sites

GO lists were downloaded using FatiGO (http://fatigo.bioinfo.cipf.es/) from Babelomics [23]. Each protein was classified with respect to its cellular component, biological

process, and molecular function using GO annotation. When no GO annotation was available, proteins were annotated manually based on literature searches and closely related homologs. The NetPhos 2.0 server (http://www.cbs.dtu.dk/services/NetPhos/) [24] was used to predict the number of potential protein phosphorylation sites. Further database searches within PhosphoBase 6.0 (http://phospho.elm.eu.org/) [25] were also carried out to confirm the phosphorylation of identified proteins and gain information about the exact position of known phosphorylated instances. Scansite 2.0 (http://scansite.mit.edu) [26] was used to search for potential Ser/Thr- or Tyr-kinases responsible for the phosphorylation by searching for specific recognized motifs within the phosphorylated protein. XplorMed (http://www.ogic.ca/projects/xplormed/) [27] was used to search the literature for signal transduction pathway reconstruction.

2.9 SNP-array analysis and correlation

DNA was extracted by column-purification methods (AllPrep columns, Quiagen). The Affymetrix GeneChip® Mapping Assay was used according to the manufacturer's instructions. Briefly, 250 ng of genomic DNA *per* array were used for restriction enzyme (Xba I) digestion and ligation of sequences onto the DNA adaptor fragments. PCR products were then fragmented, end-labeled, and hybridized to a GeneChip array following the manufacturer's instructions. Raw data analysis was performed using GTYPE 4.1. Copy number and LOH were analyzed using the Affymetrix Chromosome Copy Number Analysis Tool (CNAT) 4.0. For all samples, we performed unpaired analysis using the reference datasets from Affymetrix. To correlate the expression of identified phosphoproteins with putative underlying alterations in DNA copy number, we selected the data relative to probes mapping within 500 kb from each protein locus from the SNP-Chip and averaged their values. Next, we evaluated the presence of common alterations among cell lines, performing a series of one-sample *t*-tests on the assumption that the expected mean was 0 (no change) (Fig. 1 of Supporting Information).

3 Results

3.1 2-DE phosphoprotein pattern analysis in MCL cell lines

Representative 2-DE phosphoproteins maps for all of the MCL lines are shown in Fig. 1 together with the standard map showing the first 100 phosphoproteins that were most abundant. The typical high resolution 2-DE phosphoprotein pattern obtained from MAVER-1 cell line is reported in Fig. 2a, while representative 2-D maps of the un-phosphorylated fraction and phosphorylated fraction treated with alkaline phosphatase (from MAVER-1) are shown in Figs. 2b and c, respectively. Figure 3 shows the staining of IMAC and phosphatase-treated separated fractions by a phospho-stain such as Pro-Q Diamond, demonstrating enrichment of phosphoproteins in the IMAC separated fractions.

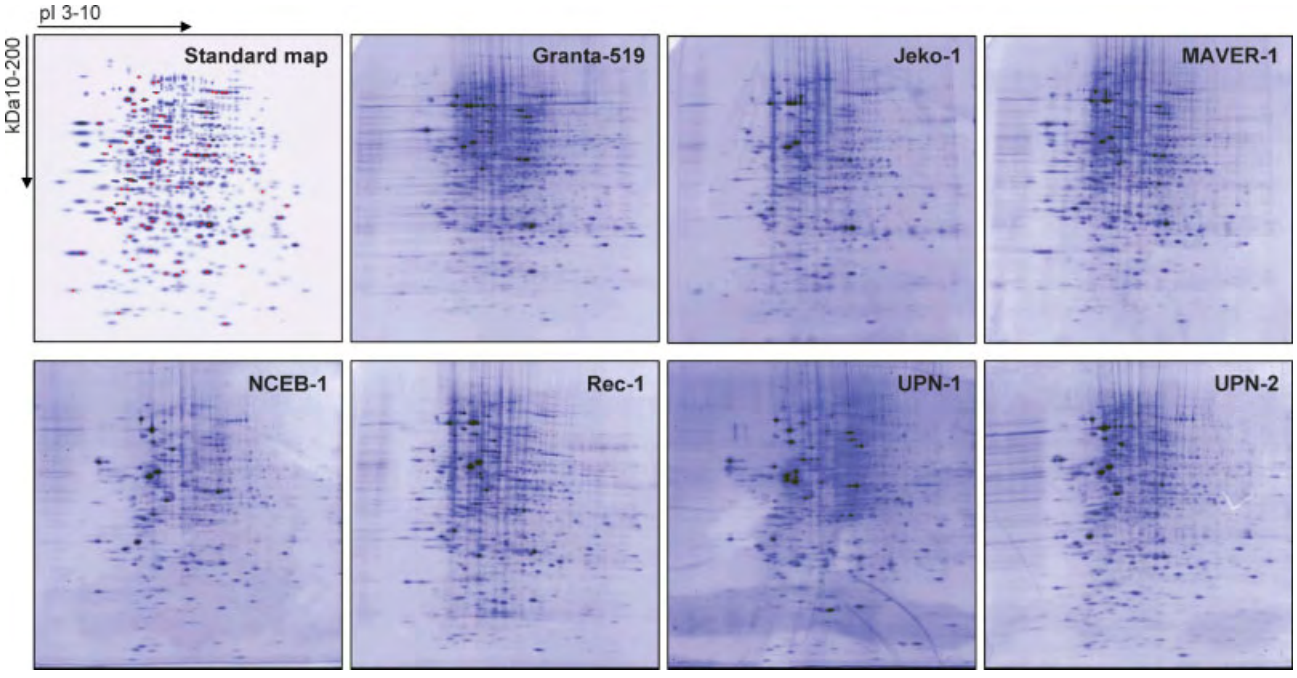

Figure 1. 2-DE phosphoproteins maps for all of the MCL lines together with the standard map showing the first 100 most abundant phosphoproteins (marked by red squares).

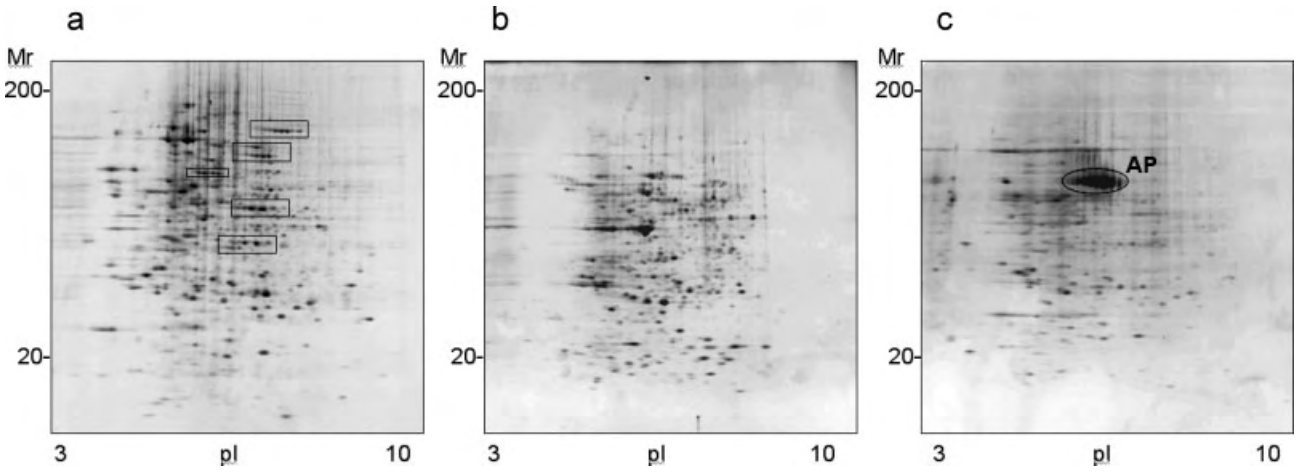

Figure 2. (a) Example of Sypro Ruby stained 2-D gel of phosphoproteins ("phosphorylation trains" are marked by rectangle); (b) un-phosphorylated fraction; and (c) the phosphorylated fraction treated with alkaline phosphatase (marked as AP) extracted from the MAVER-1 cell line.

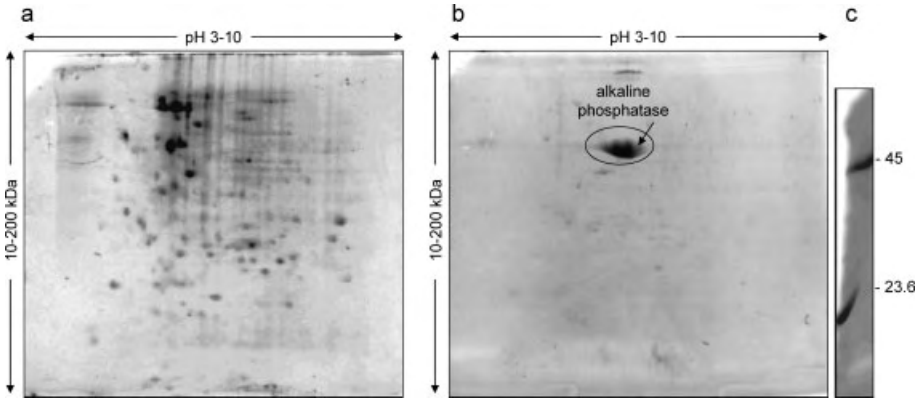

Figure 3. Phosphoprotein staining by Pro-Q Diamond. Total proteins from both (a) phosphorylated IMAC fraction and (b) phosphorylated IMAC fraction treated with protein phosphatase were separated by 2-DE gels and stained with fluorescent Pro-Q Diamond for phosphoproteins. Stained gels were scanned with a Typhoon 9600 fluorescence scanner. (c) PeppermintStickTM phosphoprotein molecular weight standards.

The 100 spots with the highest average intensity were subjected to RP-HPLC coupled with MS/MS analysis for protein identification. A total of 76 unique, highly expressed phosphoproteins were identified (Table 1 of Supporting Information).

To demonstrate the phosphorylation status of some of the identified proteins, we performed Western blot analysis on the phosphorylated and un-phosphorylated fractions. Figure 4 shows the immunoblot results obtained for some of the most relevant phosphoproteins (phospho-eIF2alpha, phospho-eukaryotic elongation factor 2 (eEF2), and phospho-cofilin) as well as a Ponceau S stain of the blots at the bottom for control of protein loading.

3.2 Bioinformatics analysis of phosphoproteins

Functional annotation of the phosphoproteins identified was carried out by categorizing the proteins into different groups based on GO terms. Figures 5a and b show a pie-chart distribution of the identified proteins, cataloged according to the cellular component or biological process (see also Table 1 of Supporting Information). Out of the 76 proteins identified, 23% were involved in RNA splicing and translation, 18% in cell proliferation, cell cycle, and apoptosis, 13% in chromatin/cytoskeleton organization and protein folding, and 9% in signal transduction. We characterized the putative phosphorylation sites of each protein by using different phosphorylation prediction programs. Using Netphos 2.0, an output score of 0.5 was used as cutoff to ensure that the site was a *bona fide* phosphorylation site. For all the identified proteins, several phosphorylation sites were predicted by Netphos 2.0. The Phosphobase 6.0 search provided the exact position of the known phosphorylation sites based on the available literature. Each confirmed phosphoprotein was further searched within Scansite 2.0, with high and medium stringency, to predict the kinase and phosphorylation-dependent binding motifs. Potential binding sites for the Erk D-domain and casein kinase 2 were most commonly predicted; casein kinase 1, PKC epsilon, and zeta sites were also found frequently (see Table 1 of Supporting Information).

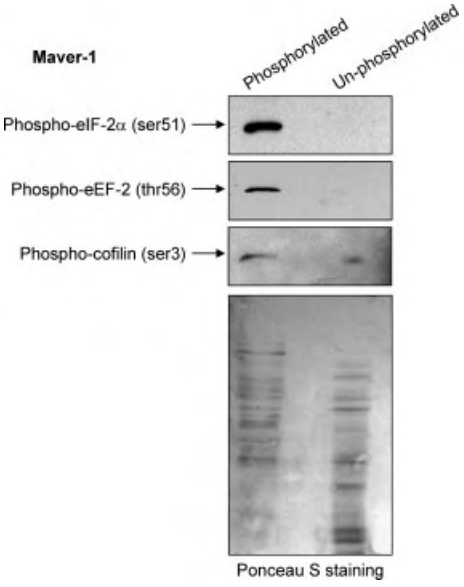

Figure 4. Protein phosphorylation validation by immunoblot analysis. Western blot analysis for three relevant phosphoproteins (Phospho-eIF2alpha, Phospho-eEF2, and Phospho-cofilin) demonstrating the enrichment of phosphorylated proteins by IMAC. A Ponceau S stain of the blots is shown at the bottom for control of protein loading.

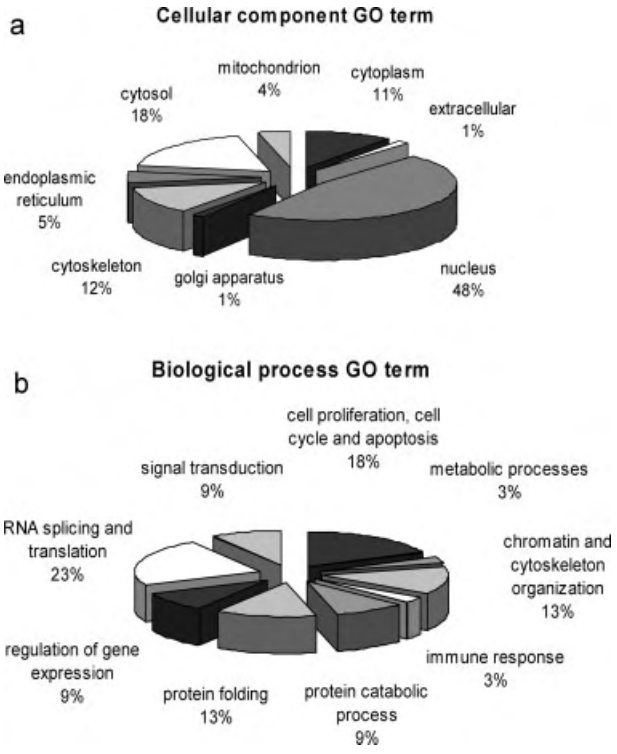

Figure 5. Distribution of identified proteins according to the (a) cellular component or the (b) biological process in which they are involved. Assignments were made on the basis of information provided by GO lists downloaded from FatiGO (http://fatigo.bioinfo.cipf.es/).

Figure 6 shows, as a simplified network, the relationships between some of the identified phosphoproteins and their involvement in some relevant signal transduction pathways.

3.3 Correlation of core phosphoproteins with SNP-Chip analysis

To correlate the expression of phosphoproteins with possible changes in DNA copy number, we extracted the information relative to the identified proteins from the SNP-Chip profiles of MCL lines. Out of 76 unique proteins, 62 had a corresponding set of probes on the SNP-Chip. The presence of significant DNA copy number alterations, among the seven cell lines investigated, was evaluated with a *t*-test, and proteins were subsequently ranked based on significance (see Fig. 1 of Supporting Information). A graphical overview of the copy number alterations is shown in Fig. 7, while a complete list is provided in Table 2 of Supporting Information. There was a trend for most highly correlated proteins to have copy-number gains, although some proteins (*e.g.*, MSN, RPSA, and RLP0) showed an inverse correlation.

4 Discussion

In this study, we were able to improve our understanding of the signal transduction pathways involved in MCL by analyzing the most expressed phosphoproteins in seven different MCL cell lines. IMAC (Qiagen) had already been successfully used to investigate phosphoproteins from plants [28], yeast [29], and mammalian cells [30]. The effectiveness of this method was further confirmed by comparing the two protein fractions. The "phosphorylation trains" that characterized the 2-D map of the phosphoprotein fraction (highlighted in Fig. 2a) are virtually absent both in the proteome profiles of the unphosphorylated fraction and in the phosphorylated fraction treated with alkaline phosphatase (Figs. 2b and c). Moreover, the enrichment of phosphoproteins by IMAC was further confirmed by Pro-Q Diamond spot staining. Using this phospho-specific stain, a large number of spots were detected in the IMAC-separated fraction; in contrast, very few spots were detectable in the phosphatase-treated sample (Fig. 3). The phosphorylation status of some of the identified proteins (Fig. 4) was demonstrated by the presence of specific bands only in the phosphorylated fractions after the Western blot analysis.

4.1 Candidate kinase identification

Bioinformatics was used to obtain information on the phosphorylation sites and the presence of kinase-specific motifs for the identified phosphoproteins (Table 1 of Supporting Information). As mentioned above, most of the identified proteins contained the typical MAPK1 and CK2 phosphorylation motifs, although casein kinase 1 and novel or atypical PCK members (epsilon and zeta) were also commonly found.

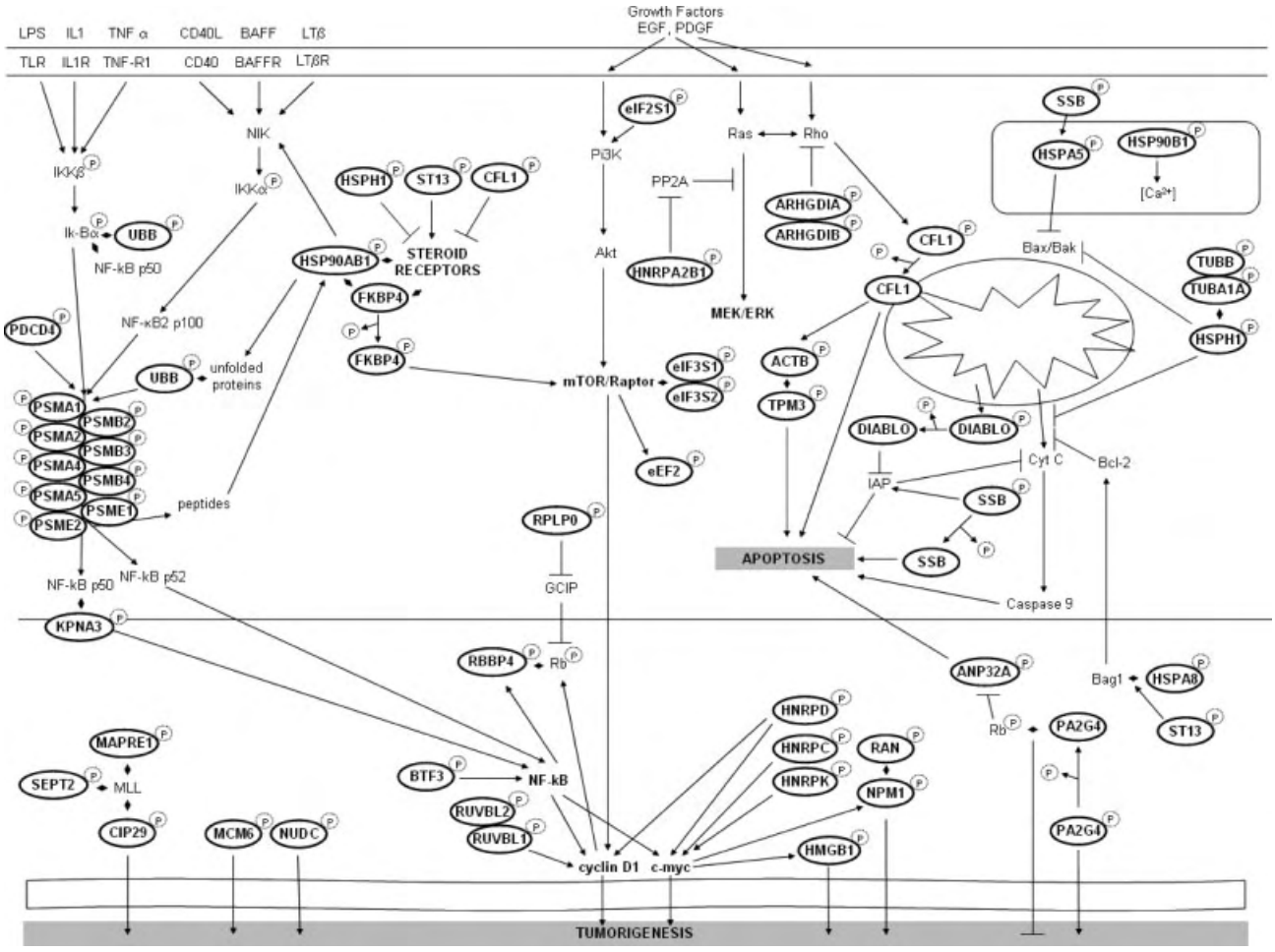

Figure 6. Chart illustrating the relationships between some of the identified phosphoproteins and their involvement in tumorigenesis-related signal transduction pathways (such as NF-κB, mTOR, mitochondrial signaling). Proteins are indicated by their HUGO gene name.

MAPK1 is involved in a wide variety of cellular processes such as proliferation, differentiation, and transcription regulation. The activation of this kinase requires its phosphorylation by upstream kinases. Interestingly, it has been reported that the constitutive activation of MAPK1 plays a key role in B-cell lymphoproliferative disorders [31], but further validation is still lacking.

CK2 is a ubiquitous serine–threonine kinase that regulates many relevant biologic processes, and is considered an oncogene [32]. CK2 is known to interact with p27 [33], and in yeast CK2 depletion blocks cell cycle by an increase in p27 homolog Sic1 [34]. Since p27 degradation is a known event in MCL pathogenesis, whose effect is to relieve its inhibition on cyclin/Cdk complexes, it is tempting to speculate that it might be favored by CK2 activity. Interestingly, CK2 also interacts with CD5 [35, 36], which is typically expressed in MCL. It has been demonstrated that CK2 is involved in the pathophysiology of multiple myeloma, suggesting that it might play a crucial role in controlling survival and sensitivity to chemotherapeutics of malignant plasma cells [37].

Novel members of the PKC family are also potentially interesting therapeutic targets [38]. PKCzeta can activate both NF-κB and Stat6 pathways [39]. Of interest, PKCzeta has been found to be active in MCL cells [40], and PKCepsilon has been shown to be a very potent inducer of Cyclin D1 expression [41]

We were able to construct a network showing the involvement of the identified phosphoproteins in some important signal transduction pathways, based on the published literature. Hereafter, some of the phosphoproteins belonging to signal transduction pathways that are relevant to tumorigenesis (including NF-κB, mTOR, and mitochondrial signaling) are briefly discussed.

4.2 NF-κB signaling related phosphoproteins

NF-κB is a transcription factor present in all cell types that is involved in cellular responses to stimuli such as stress, cytokines, free radicals, UV irradiation, and bacterial or viral antigens [42]. NF-κB is responsible for increased prolifera-

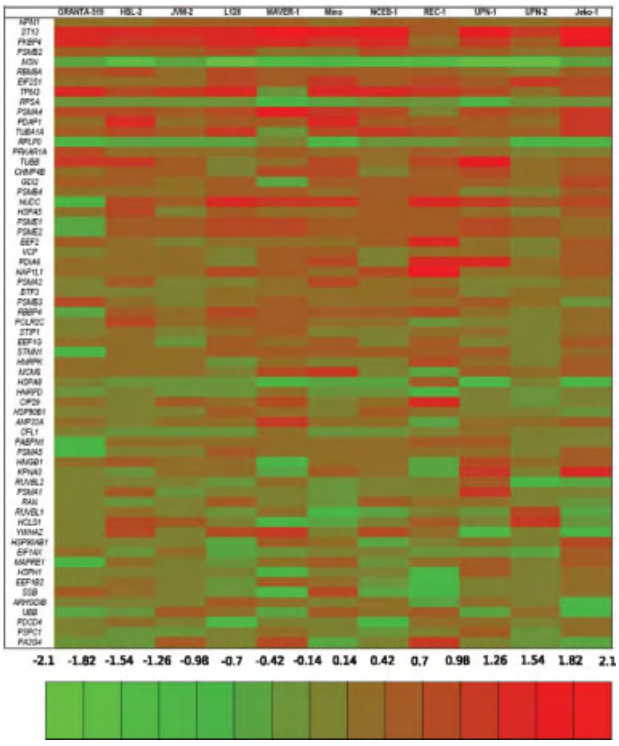

Figure 7. Overview of copy number changes in MCL cell lines in 62 loci corresponding to identified proteins. Proteins are ranked from top to bottom based on the p value of t-test (lower on top). Colors indicate copy number alteration (red: gain; green: loss). For details, see Table 2 of Supporting Information.

tion of cells and tumor growth by transcriptional activation of anti-apoptotic proteins. The nuclear translocation of NF-κB is controlled by two main pathways, namely the classical and the alternative. The classical NF-κB pathway involves the proteasome degradation of the inhibitory molecule IkBα, while the alternative NF-κB pathway induces p100 proteasome processing and p52 generation through the activation of NF-κB-inducing kinase (NIK). Gene expression profiling studies have shown that the transcripts involved in NF-κB signaling pathways are overexpressed in MCL [43] and that its inhibition induces death in MCL cells [44, 45]. Interestingly, both cyclin D1 and Bcl-2 are regulated by NF-κB [46, 47].

We found that several proteasome subunits (PSMA1, PSMA2, PSMA4, PSMA5, PSMB2, PSMB3, and PSMB4), the proteasome activator PA28 (PSME1, PSME2), and ubiquitin (UBB) are phosphorylated and highly expressed in MCL cell lines. Our results are consistent with a new MCL therapeutic strategy based on proteasome inhibition, recently approved by the FDA [48].

Among the phosphoproteins involved in the classical NF-κB pathway, we identified Importin alpha3 (KPNA3), which contributes to NF-κB nuclear transfer [49]. We also identified hsp90 (HSP90AB1), a protein involved in the alternative NF-κB pathway. This protein has been recently shown to be upregulated in MCL by antibody-microarray technology [50]. Accordingly, it has been demonstrated that HSP90AB1 protects NIK from autophagy-mediated degradation [51]. It is of interest to note that 17-allylamino-17-demethoxy-geldanamycin (17-AAG), an hsp90 inhibitor, induced G0/1 cell cycle arrest, and cell death in MCL cell lines [52]. We also identified basic transcription factor 3 (BTF3), which acts as a transcription factor and a modulator of apoptosis. This is in agreement with other findings showing that BTF3 regulates the transcription of NF-κB (and its target genes) [53]. We also found, as a phosphorylated protein, the retinoblastoma binding protein 4 (RBBP4) which is transcriptionally controlled by NF-κB [54] and belongs to the HDAC complex that binds to the tumor suppressor retinoblastoma protein Rb.

4.3 mTOR signaling-related phosphoproteins

The mTOR is a large and highly conserved kinase that integrates growth factor stimulation, energy, and nutrient availability to modulate translation of proteins responsible for cellular growth and proliferation [55]. mTOR inhibitors have demonstrated efficacy against lymphoma cells, and in particular in MCL [56]; moreover, the phosphoinositide-3 kinase (PI3K)-AKT pathway has been identified as a key player in MCL by gene expression profiling [57] as well as Western blotting and biochemical studies [58].

Among the phosphorylated proteins, we identified eukaryotic translation initiation factor 2 (EIF2S1). Previous data provided evidence that EIF2S1 phosphorylation is involved in the induction of PI3K that activates mTOR signaling [59]. We also identified eukaryotic translation initiation factor 3 (eIF3), (eIF3S1 and eIF3S2) as a phosphorylated protein. eIF3 binds to the 40S ribosomal subunit and interacts with other initiation factors and mRNA, regulating the rate of translation initiation. Recently, it has been demonstrated that eIF3 also acts as a scaffold for the mTOR/raptor complex and its substrate (S6 kinase1), allowing the coordination of protein translation [60]. Interestingly, phosphorylation of eIF3 was found to be dependent on mTOR kinase activity [61, 62]. In addition, increased levels of the eIF3 have been detected in a wide variety of human tumors and have been suggested to be prognostic for poor clinical outcome [63].

eEF2 was also identified. Along these lines, it has been reported that mTOR signaling also regulates the translation elongation process through eEF2 phosphorylation [64]. eEF2 is a GTP-binding protein that mediates the translocation step of elongation. When phosphorylated, eEF2 loses its ability to bind to ribosomes and is thus inactivated.

FK506 binding protein 4 (FKBP4, also known as FKBP52) was one of the phosphorylated proteins correlated with copy number gains. FKBP4 is another potential regulator of mTOR, as homologs (like FKBP12 and FKBP51) have been recently demonstrated to be important players in mTOR signaling [65].

4.4 Mitochondrial signaling-related phosphoproteins

Mitochondria plays important roles in cellular energy metabolism, free radical generation, and apoptosis. It is well known that defects in mitochondrial signaling contribute to the development and progression of cancer. Among the phosphoproteins related to the mitochondrial signaling, cofilin (CFL1) was highly expressed and phosphorylated in MCL. It is interesting to note that the phosphorylation of CFL1 negatively regulates its function on actin by inhibiting the depolymerization activities [66]. More importantly, it has been reported that p21 is involved in inhibition of CFL1 phosphorylation [67], and p21 is indeed downregulated in aggressive variants of MCL [68]. Accordingly, unphosphorylated CFL1 can accumulate in mitochondria and induce apoptosis [69, 70]. These findings help to explain the high level of phospho-cofilin detected in MCL.

Heat shock 110 kDa protein (HSPH1) was also identified, which is overexpressed in a variety of human tumors [71]. In HeLa cells, it suppresses apoptosis by inhibiting the translocation of Bax to mitochondria [72]. Remarkably, HSPH1 is bound directly to tubulin alpha (TUBA1A) [73], which is another phosphoprotein that we found to be highly expressed in MCL. The activation of Bax is also suppressed by heat shock protein 5 (HSPA5) 70 kDa, which was also highly expressed in MCL cell lines. Recently, it has been demonstrated that HSPA5 prevents apoptosis by inhibiting cytochrome *c* release from mitochondria [74]. It is interesting to note that the IRES-dependent translation of the anti-apoptotic HSPA5 is enhanced by autoantigen-La (SSB) [75]. SSB enhances internal ribosome entry site (IRES)-dependent translation of X-linked inhibitor of apoptosis protein (XIAP) [76]. SSB is a key protein in RNA biogenesis, and is dephosphorylated and cleaved during apoptosis [77]. SSB was among the phosphorylated proteins which are highly expressed in MCL, in agreement with a recent report exploring MCL proteome by antibody microarrays [50].

4.5 Correlation with SNP-chip analysis

While thorough analysis of all positively correlated proteins is beyond the scope of the present investigation, a few comments are merited. It has already been mentioned that FKBP4 (located on 12p13) is involved in mTOR signaling and displayed recurrent copy number gains in MCL cell lines. Another protein showing strong correlation with copy number is ST13 (also known as HIP, located on 22q13), a molecule showing similarities to BID, and probably involved in the regulation of apoptosis through interaction with BAG-1 and Hsp70 [78]. Interestingly, both these molecules appear to be involved in regulation of mitochondrial apoptosis, a pathway not widely explored in MCL. However, the role of these and other molecules deserve further investigation by independent techniques, and will constitute the core of future efforts.

5 Concluding remarks

Proteome analysis of phosphoproteins is a challenging task, and using the current approach we were able to obtain insights into the "core phosphoproteome" of MCL by analysis of seven different cell lines. We also attempted to focus on correlations between copy number gains and the most abundant phosphoproteins. Herein, we identified several signal transduction pathways involved in the biology of MCL cell lines. Some of these pathways (NF-κB and PI3K-AKT-mTOR) were recently demonstrated as key players in MCL pathogenesis, and we provide a completely independent validation of previous studies by proteomic techniques. Other pathways, such as mitochondrial signaling, are novel to MCL pathobiology. Many of the identified proteins are substrates of specific kinases whose role in MCL pathogenesis is only beginning to be understood. The importance of all of these candidate proteins will require further validation *in vitro* and *in vivo*; nevertheless, they may be promising therapeutic targets for MCL.

We thank Enrico Domenici and Lucia Carboni of the GSK Verona for sharing the spot cutter facility. This work was supported in part by grants from AIRC (to M. C.), CIB Consorzio Interuniversitario Biotecnologie and Centro Nazionale Sangue (ISS) (to L. Z.), and Interdisciplinary Center for Clinical Research (IZKF) Würzburg, Germany (to A. R. and E. H.).

The authors have declared no conflict of interest.

6 References

[1] Swerdlow, S. H., Berger, F., Isaacson, P. I., Muller-Hermelink, H. K. Mantle cell lymphoma, in: Jaffe, E. S., Harris, N. L., Stein, H., Vardiman, J. (Eds.), *Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues*, IARC Press, Lyon 2001, pp. 168–170.

[2] Jares, P., Colomer, D., Campo, E., Genetic and molecular pathogenesis of mantle cell lymphoma: Perspectives for new targeted therapeutics. *Nat. Rev. Cancer* 2007, *7*, 750–762.

[3] Rosenwald, A., Wright, G., Wiestner, A., Chan, W. C. *et al.*, The proliferation gene expression signature is a quantitative integrator of oncogenic events that predicts survival in mantle cell lymphoma. *Cancer Cell* 2003, *3*, 185–197.

[4] Antonucci, F., Chilosi, M., Parolini, C., Hamdan, M. *et al.*, Two-dimensional molecular profiling of mantle cell lymphoma. *Electrophoresis* 2003, *24*, 2376–2385.

[5] Antonucci, F., Chilosi, M., Santacatterina, M., Herbert, B., Righetti, P. G., Proteomics and immunomapping of reactive lymph-node and lymphoma. *Electrophoresis* 2002, *23*, 356–362.

[6] McLachlin, D. T., Chait, B. T., Analysis of phosphorylated proteins and peptides by mass spectrometry. *Curr. Opin. Chem. Biol.* 2001, *5*, 591–602.

[7] Mann, M., Jensen, O. N., Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* 2003, *21*, 255–261.

[8] Adamczyk, M., Gebler, J. C., Wu, J., Selective analysis of phosphopeptides within a protein mixture by chemical modification, reversible biotinylation and mass spectrometry. *Rapid Commun. Mass Spectrom.* 2001, *15*, 1481–1488.

[9] Goshe, M. B., Conrads, T. P., Panisko, E. A., Angell, N. H. *et al.*, Phosphoprotein isotope-coded affinity tag approach for isolating and quantitating phosphopeptides in proteome-wide analyses. *Anal. Chem.* 2001, *73*, 2578–2586.

[10] McLachlin, D. T., Chait, B. T., Improved beta-elimination-based affinity purification strategy for enrichment of phosphopeptides. *Anal. Chem.* 2003, *75*, 6826–6836.

[11] Oda, Y., Nagasu, T., Chait, B. T., Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nat. Biotechnol.* 2001, *19*, 379–382.

[12] Qian, W. J., Goshe, M. B., Camp, D. G., II, Yu, L. R. *et al.*, Phosphoprotein isotope-coded solid-phase tag approach for enrichment and quantitative analysis of phosphopeptides from complex mixtures. *Anal. Chem.* 2003, *75*, 5441–5450.

[13] Thaler, F., Valsasina, B., Baldi, R., Xie, J. *et al.*, A new approach to phosphoserine and phosphothreonine analysis in peptides and proteins: Chemical modification, enrichment via solid-phase reversible binding, and analysis by mass spectrometry. *Anal. Bioanal. Chem.* 2003, *376*, 366–373.

[14] Andersson, L., Porath, J., Isolation of phosphoproteins by immobilized metal (Fe3+) affinity chromatography. *Anal. Biochem.* 1986, *154*, 250–254.

[15] Posewitz, M. C., Tempst, P., Immobilized gallium(III) affinity chromatography of phosphopeptides. *Anal. Chem.* 1999, *71*, 2883–2892.

[16] Zamo, A., Ott, G., Katzenberger, T., Adam, P. *et al.*, Establishment of the MAVER-1 cell line, a model for leukemic and aggressive mantle cell lymphoma. *Haematologica* 2006, *91*, 40–47.

[17] Jadayel, D. M., Lukas, J., Nacheva, E., Bartkova, J. *et al.*, Potential role for concurrent abnormalities of the cyclin D1, p16CDKN2 and p15CDKN2B genes in certain B cell non-Hodgkin's lymphomas: Functional studies in a cell line (Granta 519). *Leukemia* 1997, *11*, 64–72.

[18] Raynaud, S. D., Bekri, S., Leroux, D., Grosgeorge, J. *et al.*, Expanded range of 11q13 breakpoints with differing patterns of cyclin D1 expression in B-cell malignancies. *Genes, Chromosomes Cancer* 1993, *8*, 80–87.

[19] Jeon, H. J., Kim, C. W., Yoshino, T., Akagi, T., Establishment and characterization of a mantle cell lymphoma cell line. *Br. J. Haematol.* 1998, *102*, 1323–1326.

[20] M'Kacher, R., Bennaceur, A., Farace, F., Lauge, A. *et al.*, Multiple molecular mechanisms contribute to radiation sensitivity in mantle cell lymphoma. *Oncogene* 2003, *22*, 7905–7912.

[21] Saltman, D. L., Cachia, P. G., Dewar, A. E., Ross, F. M. *et al.*, Characterization of a new non-Hodgkin's lymphoma cell line (NCEB-1) with a chromosomal (11:14) translocation [t(11:14)(q13;q32)]. *Blood* 1988, *72*, 2026–2030.

[22] Cecconi, D., Mion, S., Astner, H., Domenici, E. *et al.*, Proteomic analysis of rat cortical neurons after fluoxetine treatment. *Brain Res.* 2007, *1135*, 41–51.

[23] Al-Shahrour, F., Minguez, P., Tarraga, J., Montaner, D. *et al.*, BABELOMICS: A systems biology perspective in the functional annotation of genome-scale experiments. *Nucleic Acids Res.* 2006, *34*,W472–W476.

[24] Blom, N., Gammeltoft, S., Brunak, S., Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J. Mol. Biol.* 1999, *294*, 1351–1362.

[25] Blom, N., Kreegipuu, A., Brunak, S., PhosphoBase: A database of phosphorylation sites. *Nucleic Acids Res.* 1998, *26*, 382–386.

[26] Obenauer, J. C., Cantley, L. C., Yaffe, M. B., Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucleic Acids Res.* 2003, *31*, 3635–3641.

[27] Perez-Iratxeta, C., Bork, P., Andrade, M. A., XplorMed: A tool for exploring MEDLINE abstracts. *Trends Biochem. Sci.* 2001, *26*, 573–575.

[28] Chung, D. W., Pruzinska, A., Hortensteiner, S., Ort, D. R., The role of pheophorbide a oxygenase expression and activity in the canola green seed problem. *Plant Physiol.* 2006, *142*, 88–97.

[29] Makrantoni, V., Antrobus, R., Botting, C. H., Coote, P. J., Rapid enrichment and analysis of yeast phosphoproteins using affinity chromatography, 2-D PAGE and peptide mass fingerprinting. *Yeast* 2005, *22*, 401–414.

[30] Chew, C. S., Okamoto, C. T., Chen, X., Thomas, R., Drebrin E2 is differentially expressed and phosphorylated in parietal cells in the gastric mucosa. *Am. J. Physiol. Gastrointest. Liver Physiol.* 2005, *289*, G320–331.

[31] Ogasawara, T., Yasuyama, M., Kawauchi, K., Constitutive activation of extracellular signal-regulated kinase and p38 mitogen-activated protein kinase in B-cell lymphoproliferative disorders. *Int. J. Hematol.* 2003, *77*, 364–370.

[32] Seldin, D. C., Landesman-Bollag, E., Farago, M., Currier, N. *et al.*, CK2 as a positive regulator of Wnt signalling and tumourigenesis. *Mol. Cell. Biochem.* 2005, *274*, 63–67.

[33] Tapia, J. C., Bolanos-Garcia, V. M., Sayed, M., Allende, C. C., Allende, J. E., Cell cycle regulatory protein p27KIP1 is a substrate and interacts with the protein kinase CK2. *J. Cell. Biochem.* 2004, *91*, 865–879.

[34] Tripodi, F., Zinzalla, V., Vanoni, M., Alberghina, L., Coccetti, P., In CK2 inactivated cells the cyclin dependent kinase inhibitor Sic1 is involved in cell-cycle arrest before the onset of S phase. *Biochem. Biophys. Res. Commun.* 2007, *359*, 921–927.

[35] Raman, C., Kimberly, R. P., Differential CD5-dependent regulation of CD5-associated CK2 activity in mature and immature T cells: Implication on TCR/CD3-mediated activation. *J. Immunol.* 1998, *161*, 5817–5820.

[36] Raman, C., Kuo, A., Deshane, J., Litchfield, D. W., Kimberly, R. P., Regulation of casein kinase 2 by direct interaction with cell surface receptor CD5. *J. Biol. Chem.* 1998, *273*, 19183–19189.

[37] Piazza, F. A., Ruzzene, M., Gurrieri, C., Montini, B. *et al.*, Multiple myeloma cell survival relies on high activity of protein kinase CK2. *Blood* 2006, *108*, 1698–1707.

[38] Mackay, H. J., Twelves, C. J., Targeting the protein kinase C family: Are we there yet? *Nat. Rev. Cancer* 2007, *7*, 554–562.

[39] Moscat, J., Rennert, P., Diaz-Meco, M. T., PKCzeta at the crossroad of NF-kappaB and Jak1/Stat6 signaling pathways. *Cell Death Differ.* 2006, *13*, 702–711.

[40] Leseux, L., Laurent, G., Laurent, C., Rigo, M. *et al.*, PKC {zeta} mTOR pathway: A new target for rituximab therapy in follicular lymphoma. *Blood* 2008, *111*, 285–291.

[41] Soh, J. W., Weinstein, I. B., Roles of specific isoforms of protein kinase C in the transcriptional control of cyclin D1 and related genes. *J. Biol. Chem.* 2003, *278*, 34709–34716.

[42] Okamoto, T., Sanda, T., Asamitsu, K., NF-kappaB signaling and carcinogenesis. *Curr. Pharm. Des.* 2007, *13*, 447–462.

[43] Martinez, N., Camacho, F. I., Algara, P., Rodriguez, A. *et al.*, The molecular signature of mantle cell lymphoma reveals multiple signals favoring cell survival. *Cancer Res.* 2003, *63*, 8226–8232.

[44] Pham, L. V., Tamayo, A. T., Yoshimura, L. C., Lo, P., Ford, R. J., Inhibition of constitutive NF-kappaB activation in mantle cell lymphoma B cells leads to induction of cell cycle arrest and apoptosis. *J. Immunol.* 2003, *171*, 88–95.

[45] Shishodia, S., Amin, H. M., Lai, R., Aggarwal, B. B., Curcumin (diferuloylmethane) inhibits constitutive NF-kappaB activation, induces G1/S arrest, suppresses proliferation, and induces apoptosis in mantle cell lymphoma. *Biochem. Pharmacol.* 2005, *70*, 700–713.

[46] Shishodia, S., Aggarwal, B. B., Nuclear factor-kappaB activation: A question of life or death. *J. Biochem. Mol. Biol.* 2002, *35*, 28–40.

[47] Baichwal, V. R., Baeuerle, P. A., Activate NF-kappaB or die? *Curr. Biol.* 1997, *7*, R94–96.

[48] Leonard, J. P., Furman, R. R., Coleman, M., Proteasome inhibition with bortezomib: A new therapeutic strategy for non-Hodgkin's lymphoma. *Int. J. Cancer* 2006, *119*, 971–979.

[49] Fagerlund, R., Kinnunen, L., Kohler, M., Julkunen, I., Melen, K., NF-{kappa}B is transported into the nucleus by importin {alpha}3 and importin {alpha}4. *J. Biol. Chem.* 2005, *280*, 15942–15951.

[50] Ghobrial, I. M., McCormick, D. J., Kaufmann, S. H., Leontovich, A. A. *et al.*, Proteomic analysis of mantle-cell lymphoma by protein microarray. *Blood* 2005, *105*, 3722–3730.

[51] Qing, G., Yan, P., Qu, Z., Liu, H., Xiao, G., Hsp90 regulates processing of NF-[kappa]B2 p100 involving protection of NF-[kappa]B-inducing kinase (NIK) from autophagy-mediated degradation. *Cell Res.* 2007, *17*, 520–530.

[52] Georgakis, G. V., Li, Y., Younes, A., The heat shock protein 90 inhibitor 17-AAG induces cell cycle arrest and apoptosis in mantle cell lymphoma cell lines by depleting cyclin D1, Akt, Bid and activating caspase 9. *Br. J. Haematol.* 2006, *135*, 68–71.

[53] Kusumawidjaja, G., Kayed, H., Giese, N., Bauer, A. *et al.*, Basic transcription factor 3 (BTF3) regulates transcription of tumor-associated genes in pancreatic cancer cells. *Cancer Biol. Ther.* 2007, *6*, 367–376.

[54] Pacifico, F., Paolillo, M., Chiappetta, G., Crescenzi, E. *et al.*, RbAp48 is a target of nuclear factor-kappaB activity in thyroid cancer. *J. Clin. Endocrinol. Metab.* 2007, *92*, 1458–1466.

[55] Petroulakis, E., Mamane, Y., Le Bacquer, O., Shahbazian, D., Sonenberg, N., mTOR signaling: Implications for cancer and anticancer therapy [reprint in Br. J. Cancer 2007, 96, R11-5; PMID: 17393579]. *Br. J. Cancer* 2006, *94*, 195–199.

[56] Costa, L. J., Aspects of mTOR biology and the use of mTOR inhibitors in non-Hodgkin's lymphoma. *Cancer Treat. Rev.* 2007, *33*, 78–84.

[57] Rizzatti, E. G., Falcao, R. P., Panepucci, R. A., Proto-Siqueira, R. *et al.*, Gene expression profiling of mantle cell lymphoma cells reveals aberrant expression of genes from the PI3K-AKT, WNT and TGFbeta signalling pathways. *Br. J. Haematol.* 2005, *130*, 516–526.

[58] Rudelius, M., Pittaluga, S., Nishizuka, S., Pham, T. H. *et al.*, Constitutive activation of Akt contributes to the pathogenesis and survival of mantle cell lymphoma. *Blood* 2006, *108*, 1668–1676.

[59] Kazemi, S., Mounir, Z., Baltzis, D., Raven, J. F. *et al.*, A novel function of eIF2alpha kinases as inducers of the phosphoinositide-3 kinase signaling pathway. *Mol. Biol. Cell* 2007, *18*, 3635–3644.

[60] Holz, M. K., Ballif, B. A., Gygi, S. P., Blenis, J., mTOR and S6K1 mediate assembly of the translation preinitiation complex through dynamic protein interchange and ordered phosphorylation events. *Cell* 2005, *123*, 569–580.

[61] Ahlemann, M., Zeidler, R., Lang, S., Mack, B. *et al.*, Carcinoma-associated eIF3i overexpression facilitates mTOR-dependent growth transformation. *Mol. Carcinog.* 2006, *45*, 957–967.

[62] Harris, T. E., Chi, A., Shabanowitz, J., Hunt, D. F. *et al.*, mTOR-dependent stimulation of the association of eIF4G and eIF3 by insulin. *EMBO. J.* 2006, *25*, 1659–1668.

[63] Zhang, L., Pan, X., Hershey, J. W., Individual overexpression of five subunits of human translation initiation factor eIF3 promotes malignant transformation of immortal fibroblast cells. *J. Biol. Chem.* 2007, *282*, 5790–5800.

[64] Redpath, N. T., Foulstone, E. J., Proud, C. G., Regulation of translation elongation factor-2 by insulin via a rapamycin-sensitive signalling pathway. *EMBO J.* 1996, *15*, 2291–2297.

[65] Bai, X., Ma, D., Liu, A., Shen, X. *et al.*, Rheb activates mTOR by antagonizing its endogenous inhibitor, FKBP38. *Science* 2007, *318*, 977–980.

[66] Moriyama, K., Iida, K., Yahara, I., Phosphorylation of Ser-3 of cofilin regulates its essential function on actin. *Genes Cells* 1996, *1*, 73–86.

[67] Lee, S., Helfman, D. M., Cytoplasmic p21Cip1 is involved in Ras-induced inhibition of the ROCK/LIMK/cofilin pathway. *J. Biol. Chem.* 2004, *279*, 1885–1891.

[68] Pinyol, M., Hernandez, L., Cazorla, M., Balbin, M. *et al.*, Deletions and loss of expression of p16INK4a and p21Waf1 genes are associated with aggressive variants of mantle cell lymphomas. *Blood* 1997, *89*, 272–280.

[69] Chua, B. T., Volbracht, C., Tan, K. O., Li, R. *et al.*, Mitochondrial translocation of cofilin is an early step in apoptosis induction. *Nat. Cell Biol.* 2003, *5*, 1083–1089.

[70] Wang, W., Mouneimne, G., Sidani, M., Wyckoff, J. *et al.*, The activity status of cofilin is directly related to invasion, intravasation, and metastasis of mammary tumors. *J. Cell Biol.* 2006, *173*, 395–404.

[71] Kai, M., Nakatsura, T., Egami, H., Senju, S. *et al.*, Heat shock protein 105 is overexpressed in a variety of human tumors. *Oncol. Rep.* 2003, *10*, 1777–1782.

[72] Yamagishi, N., Ishihara, K., Saito, Y., Hatayama, T., Hsp105 family proteins suppress staurosporine-induced apoptosis by inhibiting the translocation of Bax to mitochondria in HeLa cells. *Exp. Cell Res.* 2006, *312*, 3215–3223.

[73] Saito, Y., Yamagishi, N., Ishihara, K., Hatayama, T., Identification of alpha-tubulin as an hsp105alpha-binding protein by the yeast two-hybrid system. *Exp. Cell Res.* 2003, *286*, 233–240.

[74] Lee, A. S., GRP78 induction in cancer: Therapeutic and prognostic implications. *Cancer Res.* 2007, *67*, 3496–3499.

[75] Kim, Y. K., Back, S. H., Rho, J., Lee, S. H., Jang, S. K., La autoantigen enhances translation of BiP mRNA. *Nucleic Acids Res.* 2001, *29*, 5009–5016.

[76] Holcik, M., Korneluk, R. G., Functional characterization of the X-linked inhibitor of apoptosis (XIAP) internal ribosome entry site element: Role of La autoantigen in XIAP translation. *Mol. Cell. Biol.* 2000, *20*, 4648–4657.

[77] Rutjes, S. A., Utz, P. J., van der Heijden, A., Broekhuis, C. *et al.*, The La (SS-B) autoantigen, a key protein in RNA biogenesis, is dephosphorylated and cleaved early during apoptosis. *Cell Death Differ.* 1999, *6*, 976–986.

[78] Caruso, J. A., Reiners, J. J., Jr., Proteolysis of HIP during apoptosis occurs within a region similar to the BID loop. *Apoptosis* 2006, *11*, 1877–1885.

Daniela Cecconi[1]
Marta Cristofoletti[1]
Alberto Milli[1]
Paolo Antonioli[1]
Sara Rinalducci[2]
Lello Zolla[2]
Giacomo Zapparoli[1]

[1]Dipartimento di Biotecnologie, University of Verona, Italy
[2]Dipartimento di Scienze Ambientali, University of Tuscia, Viterbo, Italy

Received May 20, 2008
Revised August 26, 2008
Accepted September 8, 2008

Research Article

Effect of tannic acid on *Lactobacillus plantarum* wine strain during starvation: A proteomic study

The molecular mechanisms involved in tannic-acid (TA)-mediated cell growth retardation and viability prolongation of *Lactobacillus plantarum* VP08 strain were evaluated by a proteomic analysis of starved cells grown in the presence of TA or glucose as carbon source. The tannase activity and the cell growth retardation as well as viability prolongation were confirmed by enzymatic assay and growing tests, respectively. In order to gain information about the effect triggered at the molecular level by TA, total proteins (extracted from starved cells grown in 250 mg/L TA, or 2 g/L glucose) were analyzed by a 2D-PAGE/MS approach to detect differentially expressed proteins. A total of 15 spots were found to be down-regulated and 21 up-regulated in TA-grown cells. The results indicate an overall impact of TA on proteins involved in some cellular and metabolic pathways: glycolysis, amino acid metabolism, translation and protein folding. The modulation of specific proteins correlates with the positive effect of TA on the survival of tannase-positive *L. plantarum*.

Keywords:
2-D Electrophoresis / Lactic acid bacteria / *Lactobacillus plantarum* / Tannase / Tannic acid
DOI 10.1002/elps.200800310

1 Introduction

Tannins are phenols widespread in the plant kingdom and they are mainly classified into hydrolysable and condensed tannins. Vegetable tannins are present in soil, in foodstuffs and, as a consequence, on the alimentary tracts of humans and animals [1]. Bacteria, harboring these different environments, can interact with tannins, which affect the bacteria growth [2]. For example, it has been demonstrated that tannins interfere on physiology of lactic acid bacteria (LAB) [3–5], and that tannic acid (TA), a hydrolysable tannin, has an inhibitory effect on the growth of intestinal bacteria [3]. Accordingly, we have recently demonstrated, by a proteomic approach, that TA inhibits the growth of *Lactobacillus hilgardii* (a wine spoilage bacterium) [6].

Some bacteria are able to degrade tannins by producing tannin acyl hydrolase (E.C. 3.1.1.20), commonly referred to as tannase [7, 8]. Previous investigations on tannin-degrading bacteria regarded soil species [9, 10]; afterward tannase activity has also been detected in LAB cells [11]. Several tannin-degrading LAB, such as *L. plantarum*, *L. paraplantarum* and *L. pentosus*, have been isolated from human feces and from fermented foods [12, 13], while only *L. plantarum* has been detected among the tannase-positive oenological LAB species [14]. *L. plantarum* is a versatile Gram-positive fermentative bacterium that can be found in a range of habitats, including dairy, meat and many plant fermentations. Thus, the *L. plantarum* strains are able to adapt to different environmental conditions. In fact, it has been suggested that tannase activity might confer an ecological advantage to bacteria [11, 15]. Moreover, it has been demonstrated that different bacteria strains [9, 16], and in particular *L. plantarum* [17, 18], are able to use TA as sole carbon source by inducible production of tannase.

In the last years, tannase has been the subject of many studies due to its commercial importance and complexity as catalytic molecule [8, 19]. In particular, interest on the *L. plantarum* tannase activity has been addressed for aspects of dietary of tannins-rich foods and for its application in food and beverage industrial processes [3]. Nevertheless, the information concerning the effects of tannins on the physiology of tannase-positive LAB is yet scarce. In order to

Correspondence: Dr. Daniela Cecconi, Dipartimento di Biotecnologie, Laboratorio di Proteomica, Università degli Studi di Verona, Strada le Grazie 15, 37134 Verona, Italy
E-mail: daniela.cecconi@univr.it
Fax: +39-045-8027929

Abbreviations: CFU, colony forming units; **dapD,** 2,3,4,5-tetrahydropyridine-2-carboxylate *N*-succinyltransferase; **fusA,** elongation factor G; **GO,** gene ontology; **groS,** GroES co-chaperonin; **infA,** translation initiation factor IF-1; **LAB,** lactic acid bacteria; **metC2,** cystathionine beta-lyase; **pfkA,** 6-phosphofructokinase; **pyk,** pyruvate kinase; **rpmE2,** 50S ribosomal protein L31; **rpsF,** 30S ribosomal protein S6; **rps1,** 40S ribosomal protein S1; **TA,** tannic acid; **tig,** trigger factor

elucidate the interaction between tannins and bacteria further investigations on the response of tannase-positive LAB to tannins are recommended.

In the present study, the effects of TA on a tannase-positive wine strain VP08 of *L. plantarum* were examined. Tannase activity and growth kinetics were quantified in cells cultivated in semi-synthetic medium containing TA or glucose as carbon source.

The growth kinetics analysis was performed using different TA concentrations: 100 and 250 mg/L (as enological concentrations of hydrolysable tannins that can be present in wine [20]), and 500 mg/L (as an excess amount). Glucose was used at 2 g/L in order to reproduce the residual sugar concentration present in wine after alcoholic fermentation.

Then, a comparative 2D-PAGE analysis of total proteins extracted from cells in the stationary phase (grown in 250 mg/L TA or 2 g/L glucose) was carried out. We show that TA slows down the proliferation and prolongs the viability of *L. plantarum* cells; moreover TA regulates the expression of proteins of particular interest involved in glycolysis, amino acid metabolism, translation and protein folding. The data provided insights into the possible mechanisms involved in the response of *L. plantarum* cells to tannins during starvation.

2 Material and methods

2.1 Bacterial strain and culture conditions

In this study the strain VP08 belonging to *L. plantarum* was used. This strain was isolated from red wine containing oenological tannins that were added (as winemaking practice) during the beginning of alcoholic fermentation. The strain VP08 was routinely cultivated at 28°C in MRS broth (Fluka, Seelze, Germany) pH 4.8.

2.2 Tannase assay

The enzymatic assay was carried out on whole cells and on the supernatant obtained by centrifugation of cell culture grown in MRS and after in semi-synthetic medium (composed of casamino acid 5 g/L, yeast extract 4 g/L, $MnSO_4$ 0.025 g/L, $MgSO_4$ 0.125 g/L, KCl 0.425 g/L, KH_2PO_4 0.550 mM, $CaCl_2$ 0.125 mM, pH 4.5) added with 250 mg/L TA or 2 g/L glucose. The pellet was washed twice in 2 mL of 33 mM NaH_2PO_4 buffer pH 5.0 before use of the cells for the assay. A total volume of 1 mL of the supernatant was filtered (0.22 μm pore size membranes; Millipore, Bedford, Massachusetts, USA) and concentrated 20-fold using Centricon (cut-off 5000 Da, Millipore). Tannase activity of the whole cells and supernatant was quantified by a colorimetric method as described by Nishitani and Osawa [21]. Briefly, the sample preparation was as follows: the pellet of whole cells, grown on MRS, was suspended in 1 mL of substrate medium (pH 5.0) containing 33 mM NaH_2PO_4 and 10 mM methyl gallate at final concentration to prepare a suspension whose absorbance at 600 nm was adjusted to 1 (= *ca.* 1.5×10^8 colony forming units (CFU)/mL), while the concentrated supernatant (50 μL) was suspended in 950 μL of substrate medium at a final concentration of 5% v/v. The two suspensions were incubated aerobically at 30°C for 2 h, centrifuged (8000 × *g*, 5 min), and the absorbance of supernatants was read at 450 nm.

Commercial tannase of *Aspergillus ficuum* (Sigma, St. Louis, MO, USA) was used for standard calibration. The activity was expressed in mU. By definition 1 mU of tannase hydrolyzes 1 nmol of TA *per* min at pH 5.0 at 30°C. The experiment was carried out in triplicate using independent samples. Statistical treatment of data was carried out by *t*-test and significant differences were at $p<0.05$.

2.3 Growth kinetics

The growth kinetic of the *L. plantarum* VP08 strain was tested in semi-synthetic medium, using TA as carbon source at the concentration of 100, 250 and 500 mg/L (TA 100, TA 250 and TA 500) or 2 g/L glucose.

An amount of 1 mL of overnight culture cells grown in MRS medium pH 4.8 were harvested, washed in 1 mL of NaCl 0.9% and inoculated at the concentration of 2×10^6 cell/mL in 10 mL of semi-synthetic medium contained in glass sealed tubes. The incubation was carried out at 28°C without agitation. The growth was estimated as CFU/mL by plate count on MRS agar pH 4.8, after the incubation of the plates aerobically for 3 days at 28°C. The growing test was carried out in triplicate using independent samples.

2.4 Total protein extraction

Preliminary results using different concentrations of TA and glucose allowed us to observe when cell cultures entered in the stationary phase and how long it lasted.

Thus, we have chosen 250 mg/L TA and 2 g/L glucose, and 6 days from the inoculums, as a good compromise to obtain the cells at the same physiological state (*i.e.* just before the reduction of viability) and at the same concentration during the stationary phase. Thus, a total amount of 3×10^9 cells grown in 250 mg/L TA or in 2 g/L glucose was harvested after 6 days of incubation at 28°C. The cell pellets were washed twice with 0.9% NaCl and cells were lysed, by vortexing at maximum speed for 30 min at 4°C, in 5 mL of solubilization buffer: 7 M urea (Sigma), 2 M thiourea (Sigma), 20 mM Tris (Sigma), 5 mM tributylphosphine (Sigma), 1% ampholyte pH 3–10 (Fluka, Buchs SG Switzerland), 3% CHAPS (Sigma), 1 × protease inhibitor cocktail tablet (Complete, Mini; Roche, Basel, Switzerland), added with 0.1 g micro-glass beads. The lysate were then centrifuged for 10 min at 10 000 × *g* at 4°C to remove the nucleic acids (complexed with ampholytes), as well as cellular debris,

cell membranes and unsolubilized compounds. The protein concentration was evaluated with DC Protein assay (Bio-Rad, Hercules, CA, USA) based on the Lowry method. The calibration curve was obtained by using known concentrations of BSA dissolved in distilled water.

2.5 2D-PAGE analysis of total protein extract

Total protein extracts were analyzed by 2D-PAGE on 18×20 cm slab gels (four replicas for each sample) using 17 cm immobilized 3–10 pH gradient (IPG, Bio-Rad) and 8–18% polyacrylamide gradient. IPG strips were rehydrated for 8 h with 450 μL of 2-D solubilizing solution (7 M urea, 2 M thiourea, 3% CHAPS and 20 mM Tris) containing 1.5 mg/mL of total protein from *L. plantarum* cells. IEF was carried out with a Protean IEF Cell (Bio-Rad), with a low initial linear voltage (from 0 to 1000 V over 12 h) and then by applying a final rapid voltage ramp up to 10000 V with a limiting current of 50 μA/strip. The total product time-voltage applied was 70 000 V h for each strip and the temperature was set at 20°C. For the second dimension, the IPG strips were equilibrated for 26 min by rocking in a solution of 6 M urea, 2% SDS, 20% glycerol, 375 mM Tris-HCl, pH 8.8. The IPG strips were then laid on gradient SDS-PAGE with 0.8% agarose in Tris/glycine/SDS running buffer (192 mM glycine, 0.1% SDS and Tris to pH 8.3) containing traces of bromophenol-Blue tracking dye. The second dimension was performed in a Protean Plus Dodeca cell (Bio-Rad) with Tris/glycine/SDS running buffer. The electrophoresis was conducted with continuous cooling (at 18°C) and mixing by setting a current of 40 mA for each gel for 3 min, then 2 mA/gel for 1 h, and 20 mA/gel until the tracking dye, bromophenol-Blue, reached the anodic end of the gels. The protein zones were finally revealed with Sypro Ruby stain (Bio-Rad). Gels were incubated in a fixing solution containing 40% ethanol and 10% acetic acid for 30 min followed by overnight staining in a ready-to-use Sypro Ruby solution. Destaining was performed in 10% methanol and 7% acetic acid for 1 h, followed by a rinse of at least 3 h in pure water.

2.6 Protein pattern differential analysis

Gels were scanned using a Bio-Rad VersaDoc 1000 imaging system. 2-D gel analysis was performed by PDQuest software (Bio-Rad), version 7.3. Each gel was analyzed for spot detection, background subtraction and protein spot OD intensity quantification (spot quantity definition). The gel image showing the higher number of spots and the best protein pattern was chosen as a reference template, and spots in a standard gel were then matched across all gels. Spot quantity values were normalized in each gel dividing the raw quantity of each spot by the total quantity of all the spots included in the standard gel. Gels were divided into two separated groups (TA and glucose-grown cells) and, for each protein spot, the average spot quantity value and its variance coefficient in each group were determined. Data were log transformed and quantitative and a Student's *t*-test were performed in order to compare the two groups and identify sets of proteins that showed a statistically significant difference with a confidence level of 0.05 and a minimum two-fold variation.

2.7 In-gel digestion

Spots showing a statistically significant differential expression were carefully manually cut out from 2-D Sypro Ruby stained gels and subjected to in-gel trypsin digestion according to Shevchenko *et al.* [22] with minor modifications. The gel pieces were swollen in a digestion buffer containing 50 mM NH_4HCO_3 and 12.5 ng/μL of porcine trypsin (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 to the gel pieces and digestion 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.

2.8 Peptide sequencing by nano HPLC–ESI–MS/MS

Peptide mixtures were separated by using a nanoflow-HPLC system (Ultimate; Switchos; Famos; LC Packings, Amsterdam, The Netherlands, and Agilent 1200 series, Santa Clara, CA, USA). A sample volume of 10 μL was loaded by the autosampler onto a homemade 2 cm fused-silica pre-column (75 μm id; 375 μm od; Reprosil C18-AQ, 3 μm from Ammerbuch-Entringen, DE, Germany) at a flow rate of 2 μL/min. Sequential elution of peptides was accomplished by using a flow rate of 200 nL/min and a linear gradient from Solution A (2% acetonitrile; 0.1% formic acid) to 50% of Solution B (98% acetonitrile; 0.1% formic acid) in 40 min over the pre-column in-line with a homemade 10–15 cm resolving column (75 μm id; 375 μm od; Reprosil C18-AQ, 3 μm from Ammerbuch-Entringen).

Peptides were eluted directly into a high-capacity ion trap (model HCTplus and Esquire 6000, Bruker-Daltonik, Germany). Capillary voltage was 1.5–2 kV and a dry gas flow rate of 10 L/min was used with a temperature of 230°C. The scan range used was from 300 to 1800 m/z. Protein identification was performed by searching in the National Center for Biotechnology Information non-redundant database using the Mascot program (http://www.matrixscience.com). The following parameters were adopted for database searches: complete propionamide formation on cysteines and partial oxidation of methionines, peptide Mass Tolerance $\pm$ 1.2 Da, Fragment Mass Tolerance $\pm$ 0.9 Da, missed cleavages 2. For positive identification, the score of the result of $(-10 \times \mathrm{Log}(P))$ had to be over the significance threshold level ($p<0.05$).

2.9 Protein categorization

Gene ontology (GO) lists were downloaded using the tool FatiGO [23] from Babelomics (http://fatigo.bioinfo.cipf.es/), which includes a complete suite of web tools for the functional analysis of groups of genes in high-throughput experiments along with the use of information on GO terms. Each protein was classified with respect to its biological process and molecular function using GO annotation. When no GO annotation was available, proteins were annotated manually based on literature searches and closely related homologues.

3 Results

3.1 Tannase assay

By a preliminary assay carried out in whole *L. plantarum* cells grown in MRS the strain VP08 displayed tannase activity (data not shown). This activity was quantified in cells cultivated in the medium with TA and glucose as carbon source and in the correspondent supernatants. Cells grown in TA displayed higher activity than those grown in glucose ($p = 0.0192$) (Fig. 1). In the supernatants tannase activity was absent.

3.2 Growing tests

Figure 2 shows the growth curves of the *L. plantarum* VP08 strain in medium containing 100, 250 and 500 mg/L TA and 2 g/L glucose. The kinetics differed in relationship to the source of carbon and TA concentrations. In glucose, TA 100 and TA 250 cell concentration increased 2–3 days after the inoculation, while in TA 500 cell population underwent a partial mortality in the same period. Furthermore, in TA 100 and 250 the cell population remained in the stationary phases longer (about 8 days more) than in glucose. After this growth phase a constant rate of mortality was observed.

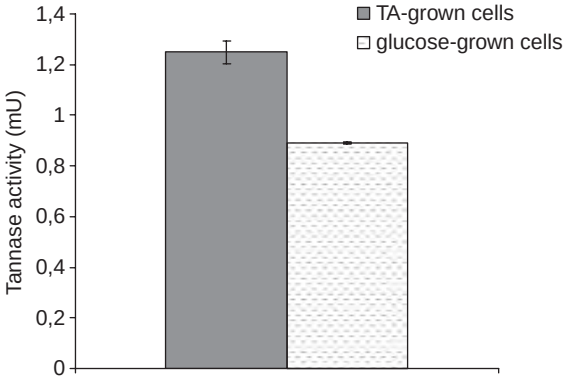

Figure 1. Tannase activity. Tannase activity of *L. plantarum* strain VP08 whole cells were grown in semi-synthetic medium containing 250 mg/L TA or 2 g/L glucose as carbon source. The tannase activity was quantified in mU *per* 1.5×10^8 CFU/mL. Values are the means of three independent experiments.

On the contrary, in TA 500 the cell number increased slowly for 12 days reaching maximal concentration as observed in TA 100 and TA 250.

3.3 L. plantarum 2-DE protein pattern analysis

To analyze the effect of TA on the growth of wine strain VP08 of *L. plantarum* the proteome profiles of cells, grown in medium supplemented with 2 g/L glucose or 250 mg/L TA, were evaluated by 2D-PAGE. Protein spots showing a statistically significant differential expression in TA-grown cells were analyzed at 6 days by means of 2-DE in the pH range of 3–10. Figure 3 shows all the 2-D maps obtained for glucose and TA-grown cells. By PDQuest analysis we detected protein spots and measured differential protein expression by analyzing the four replicas of 2-D maps obtained for each group. The average number of spots detected was 418 ± 11 and 422 ± 5 for 2-D maps of glucose and TA-grown cells, respectively. A total of 36 different spots (matched across all the replica 2-D maps) were found to be differentially expressed in cells grown in semi-synthetic medium supplemented with TA; in particular 15 spots were found to be down-regulated and 21 up-regulated. The typical high-resolution 2-DE protein pattern obtained from *L. plantarum* protein extract is reported in Fig. 4A, whereas Fig. 4B shows representative, differentially expressed spots ($p<0.05$) between glucose and TA-grown cells. Spots selected from the differential analysis were subjected to RP-HPLC–ESI–MS/MS analysis for protein identification. The unique differentially expressed proteins identified were 18. All the identifications were obtained using protein sequences of *L. plantarum* WCFS1 for protein matching. In Table 1, the successfully identified proteins corresponding to up- or down-regulated spots are shown, together with the gene names, the spot number, the identification parameters and the indication of their GO annotation (biological process and molecular function). Most of the spots were identified as proteins involved in glycolysis, amino acid metabolism, translation and protein folding. Interestingly, all the spots identified as glycolytic enzymes (glyceraldehyde 3-phospate dehydrogenase, pyruvate kinase (pyk) and 6-phosphofructokinase) increased in

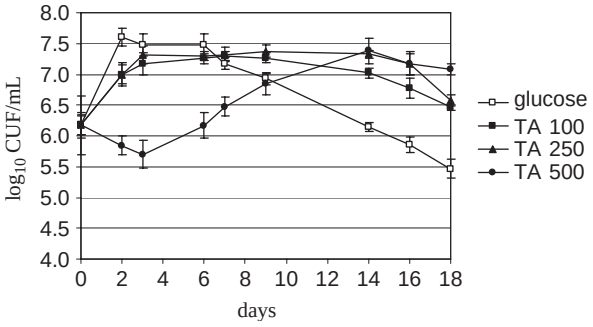

Figure 2. Growth kinetics. Growth kinetics of *L. plantarum* strain VP08 grown in semi-synthetic medium containing TA at 100, 250 and 500 mg/L or 2 g/L glucose as carbon source. Values are the means of three independent experiments.

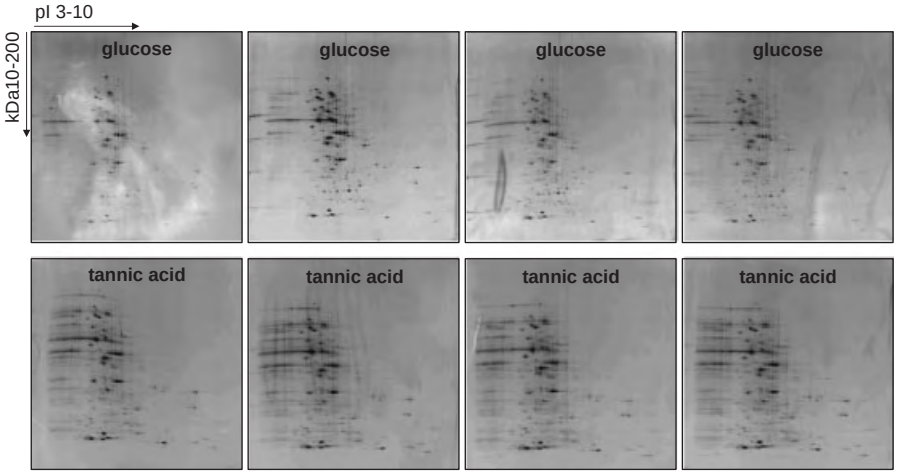

Figure 3. 2-D gel electrophoresis. *L. plantarum* 2-D maps of glucose and TA-grown cells. The proteins (about 675 μg) were first separated in a linear pH gradient of 3–10, followed by separation in an SDS-PAGE (8–18%) and Sypro Ruby fluorescent staining.

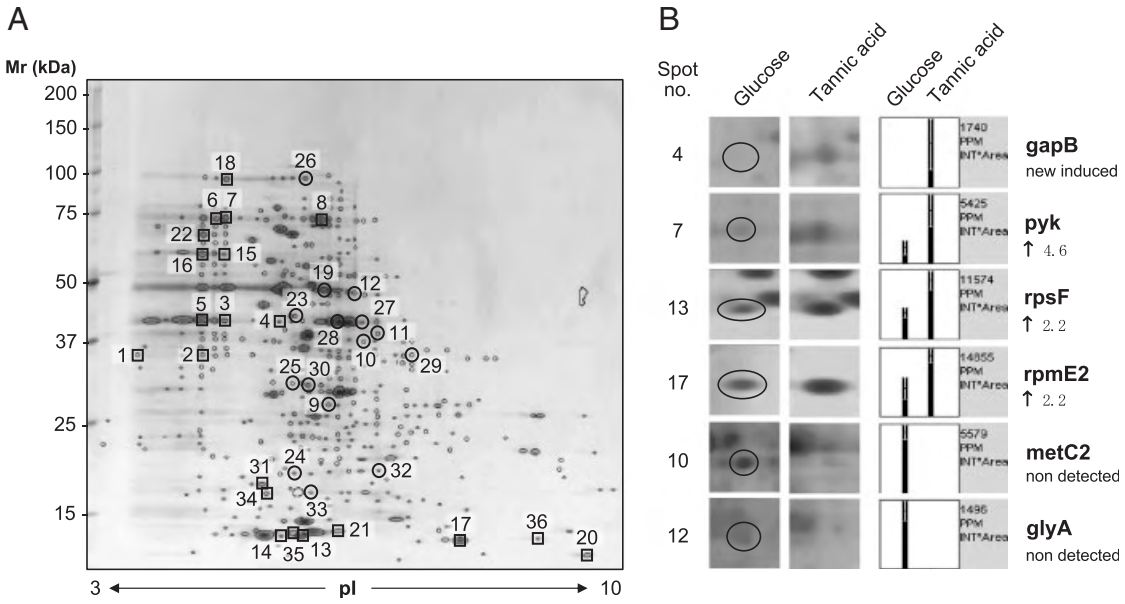

Figure 4. Protein pattern differential analysis. (A) Representative Sypro Ruby stained 2-D gels of proteins extracted from *L. plantarum* cells. The 36 differentially expressed proteins are marked by spot number as reported in Table 1, plus an open circle (to indicate decreased protein spots) or an open square (for increased protein spots). (B) PDQuest (version 7.3) output showing representative differentially expressed spots ($p<0.05$) between glucose and TA-grown cells. For each spot, an enlarged region of the respective 2-DE map containing the referred spot (highlighted by an ellipse), as well as, a histogram showing the quantitative variation in the protein expression, are shown. Bars show the average spot quantity (black dashed lines) and standard deviation (vertical white lines).

TA-grown cells. On the contrary those identified as protein related to amino acid metabolism (2,3,4,5-tetrahydropyridine-2-carboxylate *N*-succinyltransferase (dapD), cystathione β-lyase and glicine hydroxymethyltransferase) decreased in TA-grown cells. Figure 5 shows a pie chart with the distribution of the identified proteins catalogued according to the biological process in which they are involved.

4 Discussion

In this study we report that the strain VP08 of *L. plantarum* isolated from wine displays a tannase activity. This species is the only one that shows tannase activity among wine species [14]. Moreover, analysis of growth kinetics allowed us to evaluate the physiological response of cells grown in the presence of TA at enological concentrations (100 and 250 mg/L) and more (500 mg/L). Recently, we have demonstrated that, at similar experimental conditions (which mime the wine environment after alcoholic fermentation) TA produces negative effects on growth of the tannase-negative *L. hilgardii*, a wine bacterium that can co-habit with *L. plantarum* [6]. Herein, we report that growth rates of *L. plantarum* grown in the presence of TA at enological concentration (100 and 250 mg/L) are reduced, but during the stationary phase cells maintain

Table 1. Identification of the differentially expressed proteins in TA-grown cells of *L. plantarum* VP08 strain

Protein name	Gene name	Spot no. (Fig. 4)[a)]	SwissProt TrEMBL acc. #	NCBI acc. #	M_r (kDa) exp./theor.	p*I* exp./ theor.	No. of peptides identified	Mascot score[b)]	Sequence coverage (%)[c)]	Molecular function GO term	Fold of variation after TA[d)]
Glycolysis											
6-Phosphofructokinase	pfkA	1	Q88VY1	gi|28378549	27.5/34.2	4.0/5.13	3	148	12	Transferase activity	+2.7
6-Phosphofructokinase	pfkA	2	Q88VY1	gi|28378549	27.5/34.2	4.6/5.13	4	212	15	Transferase activity	New induced
Glyceraldehyde 3-phosphate dehydrogenase	gapB	3	Q88YH6	gi|28377642	33.0/36.6	4.8/5.3	6	312	22	NAD binding	+4
Glyceraldehyde 3-phosphate dehydrogenase	gapB	4	Q88YH6	gi|28377642	32.5/36.6	5.3/5.3	9	519	32	NAD binding	New induced
Glyceraldehyde 3-phosphate dehydrogenase	gapB	5	Q88YH6	gi|28377642	33.0/36.6	4.6/5.3	6	223	31	NAD binding	+3.8
Pyruvate kinase	pyk	6	Q88VY2	gi|28378548	50.0/62.9	4.7/4.99	17	1034	38	Pyk activity	+2.7
Pyruvate kinase	pyk	7	Q88VY2	gi|28378548	50.0/62.9	4.8/4.99	23	1139	51	Pyk activity	+4.6
Pyruvate kinase	pyk	8	Q88VY2	gi|28378548	60.0/62.9	5.75/4.99	23	1270	50	Pyk activity	+3.1
Amino acid metabolism											
2,3,4,5-Tetrahydropyridine-2-carboxylate *N*-succinyltransferase	dapD	9	Q88V23	gi|28378858	20.0/24.5	5.8/5.02	5	193	26	Transferase activity	−3.6
Cystathionine beta-lyase	metC2	10	Q88SB6	gi|28377181	31.5/40.8	6.15/5.52	8	351	37	Transferase activity	Non-detected
Cystathionine beta-lyase	metC2	11	Q88SB6	gi|28377181	31.5/40.8	6.25/5.52	2	123	6	Transferase activity	Non-detected
Glycine hydroxymethyltransferase	glyA	12	Q9RE02	gi|28378947	37,5/44,3	6.15/5.56	6	305	17	Pyridoxal phosphate binding	Non-detected
Translation											
30S Ribosomal protein S6	rpsF	13	Q890K2	gi|28376981	9.2/11.4	5.55/4.97	2	141	20	rRNA binding	+2.2
30S Ribosomal protein S6	rpsF	14	Q890K2	gi|28376981	9.2/11.4	5.35/4.97	2	101	20	rRNA binding	New induced
40S Ribosomal protein S1	rps1	15	P50889	gi|28378535	49.0/47.1	4.8/4.79	7	244	22	rRNA binding	+8.6
40S Ribosomal protein S1	rps1	16	P50889	gi|28378535	49.0/47.1	4.6/4.79	12	621	41	rRNA binding	+5.5
50S Ribosomal protein L31	rpmE2	17	Q88Z52	gi|28377400	9.0/9.1	7.15/6.82	2	70	29	rRNA binding	+2.2
Elongation factor G	fusA	18	Q88XY8	gi|28377830	77.5/76.9	4.8/4.81	7	370	12	GTP binding	+3.4
Elongation factor Tu	tuf	19	Q88VE0	gi|28378740	40.0/43.3	5.8/4.95	13	718	37	GTP binding	−2.1
Translation initiation factor IF-1	infA	20	Q88XW4	gi|28377854	8.2/8.2	8.4/8.04	4	174	55	RNA binding	+2.7
Protein folding											
GroES co-chaperonin	groS	21	Q88YM6	gi|28377590	9.5/10.3	5.9/4.95	3	104	42	Cpn60 binding	+2
Trigger factor	tig	22	Q88VE1	gi|28378739	56.0/49.4	4.6/4.5	13	589	36	Isomerase activity	+4.8

Table 1. Cont.

Protein name	Gene name	Spot no. (Fig. 4)[a]	SwissProt TrEMBL acc. #	NCBI acc. #	M_r (kDa) exp./theor.	p/exp./theor.	No. of peptides identified	Mascot score[b]	Sequence coverage (%)[c]	Molecular function GO term	Fold of variation after TA[d]
Transcription											
DNA-directed RNA polymerase alpha subunit	rpoA	23	Q88XW0	gi|28377858	34.5/34.9	5.5/4.8	6	236	19	DNA binding	−2
Other metabolic processes											
Hypothetical protein 1p_1747(Universal stress protein family like)		24		gi|28378427	13.5/18.1	5.5/4.74	4	243	34	ATP binding	−4.3
Malolactic enzyme	mleS	22	Q88XR8	gi|28377905	56.0/59.5	4.6/4.72	5	175	14	Malic enzyme activity	+4.8
Short-chain dehydrogenase/ oxidoreductase	yusZ	25	Q88SD1	gi|28379469	23.5/31.6	5.45/4.85	2	113	8	Oxidoreductase activity	−2.5

a) Spot numbers refer to those in Fig. 4.
b) Score is $-10^6 \log(p)$, where p is the probability that the observed match is a random event, based on the NCBI database using the MASCOT searching program as MS/MS data.
c) Amino acid sequence coverage for the identified protein.
d) Fold of variation in expression in glucose *versus* TA-grown cell: increased protein (+) and decreased protein (−).

Figure 5. Protein GO categorization. Distribution of the identified proteins according to the biological function. Assignments were made on the basis of information provided by GO lists downloaded using the tool FatiGO from Babelomics (http://fatigo.bioinfo.cipf.es/). The absolute number of proteins to which the distribution is referred is: 6 (translation), 3 (glycolysis), 3 (amino acid metabolism), 2 (protein folding), 1 (transcription) and 3 (other metabolic processes).

a higher viability than glucose-grown cells. The growth test also revealed that high concentration of tannins (500 mg/L TA) resulted initially in a strong inhibitory effect according to the previous investigation [6], with the subsequent recovering of cells viability, as a consequence of cell adaptation to tannins.

To analyze the molecular mechanisms involved in TA-mediated growth retardation and viability prolongation, we performed a proteomic analysis of starved cells grown in the presence of TA or glucose as carbon source. Our data suggest that TA does not have global effect on protein expression, but rather alters the expression of a relatively low number of proteins belonging to important cellular and metabolic pathways.

For example, here we show that 6-phosphofructokinase (pfkA), glyceraldehyde 3-phosphate dehydrogenase (gapB) and pyk are up-regulated in TA-grown cells during starvation. It is well known that exhaustion of an essential nutrient, and/or accumulation of a fermentation end product, limits the exponential phase of growth of microorganisms causing them to enter into the stationary phase [24, 25]. It has also been reported that during the early-stationary phase *L. plantarum* reduces the level of proteins involved in energy metabolic pathways [26]. The up-regulation of glycolytic enzymes that we report in this study suggests that TA-grown cells are able to obtain additional energy even during the starvation phase. It is possible to speculate that the tannase-positive strain VP08 of *L. plantarum* is able to gain supplementary energy by hydrolyzing TA. Indeed, hydrolyzable tannins, such as TA, are composed of esters of gallic acid (gallotannins) with a sugar core that is usually glucose, and are readily hydrolyzed by tannase into monomeric products [9]. The releasing of glucose by tannase, and the consequent upregulation of glycolytic enzymes (such as pfkA, gapB and pyk), could correlate with the viability

prolongation of TA-grown *L. plantarum* cells during the starvation phase.

Here, we also show a decrease of dapD, cystathionine beta-lyase (metC2) and glycine hydroxymethyltransferase (glyA) in TA-grown cells. DapD is a protein involved in meso-diaminopimelate/lysine biosynthetic pathway, metC2 in the degradation of cysteine and methionine [27], while glyA catalyzes the conversion of serine to glycine. The maintenance of an active metabolic state is common for LAB survival during the stationary phase [24, 25]. During starvation, bacteria are capable of using alternative carbon sources such as RNA, lipids, proteins peptides and amino acids. In particular, catabolism of amino acids plays an important role in the survival of *L. lactis* [28] and *L. sakei* [29] during the stationary phase. In contrast with this findings our data suggest that catabolism of amino acids does not play a pivotal role in the survival of TA-grown *L. plantarum* cells. However, the correlation between dapD, metC2 and glyA down-regulation and the prolonged viability of TA-grown cells remain to be fully clarified.

In this study, we also found that 30S ribosomal protein S6 (rpsF), 40S ribosomal protein S1 (rps1), 50S ribosomal protein L31 (rpmE2), translation initiation factor IF-1 (infA) and elongation factor G (fusA) are up-regulated in TA-grown cells. The ribosomal proteins rpsF and rpmE2 bind to 16S and 23S rRNA, respectively, rps1 maintains the mRNA in a single-stranded state to facilitate the 30S ribosome binding, infA regulates the 30S subunit/mRNA/initiator – tRNA complex formation, while fusA facilitates the translocation of the peptidyl-tRNA from the A-site to the P-site of the ribosome. It has been demonstrated that the bacteria growth rate and survival are directly correlated to the protein synthesizing capacity. For example, in *L. lactis* ribosomal proteins decreased in concentration as a function of the growth rate [30]. Here, we show up-regulation of translation related proteins, which might reflect the effect of TA on *L. plantarum* cells viability prolongation.

In this study we also report that GroES co-chaperonin (groS) and trigger factor (tig) are up-regulated in TA-grown cells. The groS is part of the chaperonin GroEL that binds, refolds and releases denatured protein, while tig protein is a peptidylprolyl isomerase that catalyzes protein folding, associates with nascent polypeptides on ribosomes and cooperates with the GroEL chaperone in promoting the degradation of unstable proteins. It has been already demonstrated that both these proteins are up-regulated in heat shock response of *L. plantarum* [24]. Our data suggest that TA-grown cells are characterized by an active folding machine, which probably assure the potential to adapt to the TA stressful conditions and survive longer during starvation. Further analysis will be required on independent samples to confirm the association of the proteins we detected with the growth and survival of *L. plantarum* in wine after alcoholic fermentation.

In conclusion, this study represented the first contribution to understand the mechanisms involved in TA-mediated cell growth retardation and viability prolongation of *L. plantarum* under starvation stress (a condition that can occur frequently in natural environment as wine). These findings might have applications in future, especially in the field of biotechnologies, for the selection or the engineering of new malolactic starters, and for the evaluation of fitness and level of adaptation of strains under starvation stress in wine.

L.Z. was financially supported by the Italian Ministry for University and Research (MIUR-PRIN 2006) and by MIPAF Genzoot 2006 Project.

The authors have declared no conflcit of interest.

5 References

[1] Haslam, E., *Practical Polyphenolics. From Structure to Molecular Recognition and Physiological Action*, Cambridge University Press, Cambrige, U.K. 1998.

[2] Scalbert, A., *Phytochemistry* 1992, *30*, 3875–3883.

[3] Chung, K.-T., Wong, T. Y., Wei, C.-I., Huang, Y.-W. Lin, Y., *Crit. Rev. Food Sci. Nutr.* 1998, *38*, 421–464.

[4] Alberto, M. R., Farìas, M. E., Manca De Nadra, M. C., *J. Food Prot.* 2002, *65*, 211–213.

[5] Renouf, V., Claisse, O., Miot-Sertier, C., Lonvaud-Funel, A., *Food Microbiol.* 2006, *23*, 136–145.

[6] Bossi, A., Rinalducci, S., Zolla, L., Antonioli, P. *et al.*, *J. Appl. Microbiol.* 2007, *102*, 787–795.

[7] Mingshu, L., Kai, Y., Qiang, H., Dongying, J., *J. Basic Microbiol.* 2006, *1*, 68–84.

[8] Aguilar, C., Rodrìguez, R., Gutiérrez-Sanchez, G., Augur, C. *et al.*, *Appl. Microbiol. Biotechnol.* 2007, *76*, 47–59.

[9] Deschamps, A. M., Otuk, G., Lebeault, J. M., *J. Ferment. Technol.* 1983, *61*, 55–59.

[10] Mondal, K. C., Banerjee, R., Pati, B. R., *Biotechnol. Lett.* 2000, *0*, 767–769.

[11] Osawa, R., Kuroiso, K., Goto, S., Shimizu, A., *Appl. Environ. Microbiol.* 2000, *66*, 3093–3097.

[12] Nishitani, Y., Sasaki, E., Fujisawa, T., Osawa, R., *Syst. Appl. Microbiol.* 2004, *27*, 109-117.

[13] Sabu, A., Augur, C., Swati, C., Pandey, A., *Process Biochem.* 2006, *41*, 575–580.

[14] Vaquero, I., Marcobal, A., Muñoz, R., *Int. J. Food Microbiol.* 2004, *96*, 199–204.

[15] Sasaki, E., Shimada, T., Osawa, R., Nishitani, Y. *et al.*, *Syst. Appl. Microbiol.* 2005, *28*, 358–365.

[16] Chowdhury, S. P., Khanna, S., Verma, S. C., Tripathi, A. K., *J. Appl. Microbiol.* 2004, *97*, 1210–1219.

[17] Ayed, L., Hamdi, M., *Biotechnol. Lett.* 2002, *24*, 1763–1765.

[18] Rodrìguez, H., De Las Rivas, B., Gòmez-Cordovéz, C., Muñoz, R., *Int. J. Food Microbiol.* 2008, *121*, 92–98.

[19] Lekha, P. K., Lonsane, B., *Adv. Appl. Microbiol.* 1997, *44*, 215–260.

[20] Waterhouse, A. L ., *Ann. NY Acad. Sci.* 2002, *957*, 21–36.

[21] Nishitani, Y., Osawa, R., *J. Microb. Methods* 2003, *54*, 281–284.

[22] Shevchenko, A., Wilm, M., Vorm, O., Mann, M., *Anal. Chem.* 1996, *68*, 850–858.

[23] Al-Shahrour, F., Minguez, P., Tárraga, J., Montaner, D. *et al.*, *Nucleic Acids Res.* 2006, *34*, 850–858.

[24] De Angelis, M., Gobbetti, M., *Proteomics*, 2004, *4*, 106–122.

[25] Ganesan, B., Stuart, M., Weimer, B. C., *Appl. Environ. Microbiol.* 2007, *73*, 2498–2512.

[26] Cohen, D. P., Renes, J., Bouwman, F. G., Zoetendal, E. G. *et al.*, *Proteomics* 2006, *6*, 6485–6493.

[27] Irmelr, S., Raboud, S., Beisert, B., Rauhut, D., Berthoud, H., *Appl. Environ. Microbiol.* 2008, *74*, 99–106.

[28] Stuart, M. R., Chou, L. S., Weimer, B. C., *Appl. Environ.Microbiol.* 1999, *65*, 665–673.

[29] Champomier Vergès, M. C., Zúniga, M., Morel-Deville, F., Pèrez-Martinez, G. *et al.*, *FEMS Microbiol. Lett.* 1999, *180*, 297–304

[30] Palmfeldt, J., Levander, F., Hahn-Hägerdal, B., James, P., *Proteomics* 2004, *4*, 3881–3898.

Ringraziamenti

Desidero ringraziare tutte le persone che hanno permesso il compimento del mio lavoro, con i loro contributi, il loro supporto e le loro idee. In particolare ringrazio il Prof. Pier Giorgio Righetti, il Prof. Lello Zolla e tutto il gruppo di Viterbo, il Prof. Franco Zunino, il Dott. Giacomo Zapparoli, la Dott.ssa Annalisa Polverari e infine, ma non certo per importanza, il Dott. Alessandro Vindigni e il Dott. Federico Odreman.

Un ringraziamento particolare va alla Dott.ssa Daniela Cecconi e alla Dott.ssa Alessandra Maria Bossi per essere riuscite a tenere in vita e dare vigore ad un laboratorio rimasto inguaribilmente "acefalo", ma non per questo meno valido, permettendomi così di compiere il mio corso. In special modo ringrazio Daniela, che mi ha costantemente e pazientemente guidato, spronato e aiutato, e che, anche nel correggermi, mi ha sempre dimostrato stima che sento di ricambiare largamente.

Un ringraziamento e un abbraccio particolari vanno a tutti i fantastici dottorandi, tesisti e tirocinanti del Laboratorio, a quelli passati come a quelli ancora presenti, che hanno contribuito, ognuno a suo modo, a creare un eccezionale gruppo di lavoro e non solo. Ringrazio soprattutto Rita, che pur dovendomi - per forza di cose - sopportare, non mi ha mai negato i suoi preziosi consigli e la sua sorridente e viva amicizia. Grazie a Davide, che con la sua innata simpatia è sempre riuscito ad alleggerirmi anche le giornate più gravose, e a Stefano, il migliore dei tesisti, per l'ilarità che mi ha regalato.

Un caloroso ringraziamento va a tutti gli amici, che durante questi anni si sono sempre fatti trovare pronti comunque e dovunque per divertirsi assieme a me e che con me hanno assaporato intensi sorsi dal calice della vita. Un pensiero particolare va al Karich, al Cobe e al Mila, i migliori compagni delle migliori trasferte. Non voglio dimenticare nemmeno Gherardo, il Gine e Il Dadde, amici di lunga data, e così Stefano C. e Alessandro C., Carlos e Michela, e via via tutti gli altri, fino a coloro con cui ho avuto il piacere di passare almeno una serata assieme.

È con il cuore che desidero ringraziare Jessica, che in tutti questi anni, strappandomi ai demoni della mia mente e conducendomi con sé tra gli dei dell'Olimpo, mi ha indefessamente accompagnato e ispirato, incoraggiato e rigenerato con vero e gioioso amore.

Voglio infine ringraziare la mia famiglia, la più bella che io potessi desiderare. Ringrazio mio padre Fernando e mia madre Rosanna, cui devo tutto ciò che sono e che non mi hanno mai negato il loro incondizionato aiuto e amorevole sostegno. E ringrazio mio fratello Corrado e mia sorella Gloria, gli amici di gran lunga più fedeli e leali che io conosca.

Di nuovo grazie a tutti, ogni volta che guardo la mia immagine riflessa nello specchio, vedo anche ognuno di voi.