

---

**METODOLOGIE      INNOVATIVE      IN**  
**PROTEOMICA CLINICA**

---

**Marianna Caterino**

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





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



---

# **METODOLOGIE INNOVATIVE IN PROTEOMICA CLINICA**

---

**Marianna Caterino**

Dottoranda: Marianna Caterino

Relatore: Prof. Margherita Ruoppolo

Coordinatore: Prof. Gennaro Marino



*A Gabriella per essere  
stata negli anni  
un modello e un sostegno*

## INDICE

<b>ABBREVIAZIONI</b>	pag 3
<b>RIASSUNTO</b>	pag 4
<b>SUMMARY</b>	pag 8
<b>Capitolo 1 INTRODUZIONE</b>	
1.1 Premesse scientifiche	pag 12
1.2 Proteomica funzionale	pag 13
1.3 Proteomica differenziale	pag 16
1.4 Spettrometria di massa e proteomica	pag 19
1.4.1 Strumentazione	pag 19
1.4.2 Strategie	pag 21
1.5 Scopo del progetto di tesi	pag 22
<b>Capitolo 2. PROCEDURE SPERIMENTALI</b>	
2.1 Estrazione di proteine di membrane da RBC	pag 23
2.2 Separazione elettroforetica di estratti proteici di membrana, colorazione e analisi d'immagine	pag 23
2.3 Analisi mediante <i>western blotting</i> di estratti proteici di membrana	pag 24
2.4 Estrazione di proteine da cellule P19	pag 24
2.5 Controllo dell'espressione di Tbx1 in cellule P19-Tbx1-TEV-PA	pag 25
2.6 Saggi di controllo: <i>labelling efficiency</i> e <i>Same same same</i>	pag 25
2.7 Marcatura covalente di estratti proteici da cellule P19 con i fluorofori CyDye DIGE	pag 26
2.8 Separazione elettroforetica di estratti proteici da cellule P19, colorazione	pag 27
2.9 Analisi d'immagine mediante il programma DyCyder Differential Analysis	pag 27
2.10 Purificazione dei ligandi delle proteine di fusione GST e GST-Rps19	pag 27
2.11 Analisi mediante <i>western blotting</i> e immunoprecipitazione dei ligandi delle proteine di fusione GST e GST-Rps19	pag 28
2.12 Analisi <i>in silico</i> degli interattori proteici di Rps19	pag 28
2.13 Separazione elettroforetica dei ligandi delle proteine di fusione GST e GST-Rps19	pag 29
2.14 Idrolisi <i>in situ</i>	pag 29
2.15 Spettrometria di massa MALDI TOF e analisi dei dati	pag 29
2.16 Spettrometria di massa <i>tandem</i> e analisi dei dati	pag 30
<b>Capitolo 3. CARATTERIZZAZIONE DELLE PROTEINE DI MEMBRANA DEL GLOBULO ROSSO IN FUNZIONE DELLA DENSITA' CELLULARE</b>	
3.1 Introduzione	pag 32
3.1.2 L'approccio proteomico allo studio delle proteine di membrana	pag 33
3.2 Risultati	pag 35
3.3 Discussione	pag 41

---

**Capitolo 4. CARATTERIZZAZIONE DEI TARGETS MOLECOLARI DEL FATTORE TRASCRIZIONALE TBX1. SINDROME DI DiGEORGE**

4.1	Introduzione	pag 43
4.2	Risultati	pag 45
4.3	Discussione	pag 56

**Capitolo 5. ANALISI DELL'INTERATTOMA DELLA PROTEINA RIBOSOMIALE S19**

5.1	Introduzione	pag 58
5.1.2	Biogenesi del ribosoma	pag 58
5.2	Risultati	pag 60
5.2.1	Identificazione dei ligandi proteici di Rps19	pag 60
5.2.2	Validazione dei dati di spettrometria di massa mediante <i>western blotting</i> e immunoprecipitazione	pag 85
5.2.3	Confronto tra l'analisi <i>in silico</i> e l'analisi sperimentale	pag 85
5.3	Discussione	pag 89

<b>CONCLUSIONI</b>	pag 91
<b>BIBLIOGRAFIA</b>	pag 92
<b>PUBBLICAZIONI E COMUNICAZIONI</b>	pag 100

## ABBREVIAZIONI

2DE	Bidimensional Electrophoresis
ASB	Amminosulfobetaine
CHAPS	3-[(3-Cholamidopropyl)Dimethylammonio]-1-Propanesulfonate Hydrate
DBA	Diamone Blackfan Anemia
DGS	DiGeorge Sindrome
DIGE	Differential Gel Electrophoresis
DMEM	Dulbecco-Modified Minimal Essential Medium
DMSO	Dimethyl Sulfoxide
DTT	Ditiotreitolo
ECL	Enhanced Chemio-Luminescence
EDTA	Ethylenediamine-Tetraacetic Acid
ESI	Electrospray Ionization
FBS	Fetal Bovine Serum
GST	Glutathione S-Trasferase
HPLC	High Performance Liquid Chromatography
HRPD	Human Protein Reference Database
HS	Hereditary Spherocytosis
IAM	Iodoacetamide
IEF	Isoelectrofocusing
IPG	Immobyline Polyacrilamide Gel
IT	Ion Trap
LC	Liquid Chromatography
MALDI	Matrix Assisted Laser Desorption Ionisation
Ms	Mass Spectrometry
NCBI	National Center For Biotechnology Information
PAGE	Polyacrilamide Gel Electrophoresis
PBS	Phosphate Buffered Saline
PMSF	Phenylmethanesulfonyl Fluoride
Q	Quadrupole
RAR	Recettore dell'Acido Retinoico
RBC	Red Blood Cell
RP	Ribosomal Protein
RPS19	Ribosomal Protein S19 gene
Rps19	Ribosomal Protein S19
SDS	Sodio Dodecil Solfato
SnoRNPS	<i>Small Nucleolar Ribonucleoprotein</i>
TAP	Tandem Affinity Purification
TBE	T-Box Binding Element
TBP	Tributylphosphine
TBX1	T-box protein 1 gene
Tbx1	T-box protein 1
TEV	Tabacco Etch Virus Protease
TOF	Time Of Flight (Tempo Di Volo)

## RIASSUNTO

La proteomica è definita come lo studio della globalità delle proteine, nella loro complessa attualità funzionale. La conoscenza della sequenza dell'intero genoma umano costituisce solo il primo stadio per la comprensione del funzionamento della macchina cellulare a livello molecolare. Gli attuali studi proteomici sono divisi in due principali aree di approfondimento: la proteomica di espressione, volta alla definizione qualitativa e quantitativa dei livelli di espressione proteica dipendenti da particolari stimoli o condizioni fisiopatologiche; e la proteomica funzionale per la definizione di complessi multiproteici. Ambedue gli approcci metodologici si basano essenzialmente su due passaggi analitici consecutivi: separazione delle proteine di un proteoma e loro successiva identificazione individuale. Il frazionamento delle miscele proteiche è realizzato mediante elettroforesi mono e/o bidimensionale su gel d'acrilammide. La successiva identificazione delle specie individuali è realizzata mediante spettrometria di massa, coadiuvata da sistemi computazionali che facilitano la ricerca in banca dati. In particolare il progetto di tesi di dottorato si è articolato nei seguenti punti:

- ✓ Caratterizzazione delle proteine di membrana del globulo rosso, in funzione della densità cellulare
- ✓ Caratterizzazione dei target molecolari del fattore trascrizionale TBX1, coinvolto nella sindrome di DiGeorge
- ✓ Studio dell'interattoma della proteina RPS19. Implicazioni nell'anemia di Diamond Blackfan

### **Caratterizzazione delle proteine di membrana del globulo rosso, in funzione della densità cellulare**

L'eritrocita manifesta le sue funzioni grazie all'interazione di tre elementi, quali la membrana eritrocitaria, l'emoglobina e le vie metaboliche che consentono di regolare il trasporto di ossigeno, proteggere l'emoglobina dai danni ossidativi e preservare la stabilità osmotica dell'ambiente cellulare. L'organizzazione strutturale e la composizione biochimica sono determinanti per la funzionalità della membrana eritroide; da esse dipendono, infatti, parametri cellulari, quali la forma, il volume cellulare medio, la resistenza osmotica e la deformabilità della cellula stessa. La membrana eritroide è costituita da un doppio strato lipidico e da proteine integrali legate ad uno scheletro proteico sottostante. Difetti nelle interazioni proteina-proteina lungo la membrana, o tra membrana e citoscheletro, o nella composizione proteica di membrana, determinano anomalie nei rapporti superficie-volume e nelle concentrazioni dei cationi monovalenti, destabilizzando il doppio strato lipidico con conseguente perdita del materiale proteico. La cellula, pertanto, assume una forma sferica che modifica proprietà della membrana, quali plasticità e deformabilità, indispensabili alla sua sopravvivenza nel microcircolo e alla sua resistenza osmotica. Le anomalie elencate costituiscono le basi molecolari di un'ampia classe di patologie, note come sferocitosi. Fenotipi differenti caratterizzano le sferocitosi: le più comuni forme della patologia presentano rispettivamente basse concentrazioni di due proteine del citoscheletro, spettina e anchirina, e di una proteina di membrana, banda 3. Una riduzione dei sintomi patologici delle sferocitosi per deficit di anchirina e spettina o per deficit di banda 3 è stata osservata in seguito alla rimozione della milza, un organo preposto alla maturazione di cellule eritroidi immature (neociti) a cellule mature (gerociti). Per poter studiare l'organizzazione strutturale, la composizione e l'evoluzione temporale della membrana eritroide sono stati confrontati i proteomi di due popolazioni cellulari: cellule eritroidi immature (neociti) ed adulte (gerociti). Queste ultime, infatti, pur derivando dal medesimo campione, differiscono per la diversa densità, deformabilità e

comportamento in processi cellulari, quali ad esempio la vescicolazione. Le due frazioni cellulari sono state separate con un gradiente di densità e le proteine di membrana sono state analizzate mediante elettroforesi bidimensionale. L'analisi d'immagine comparativa delle mappe ottenute evidenzia una serie di spots, il cui profilo di espressione differisce nei campioni analizzati. In particolare nella mappa bidimensionale della frazione dei gerociti sono presenti alcuni spots che mancano nella frazione di cellule immature. Procedure di spettrometria di massa hanno permesso di identificare alcuni spots come isoforme dell'anchirina R ed altri come prodotti proteolitici del frammento citosolico della banda 3. Tali specie proteiche sono una conseguenza del processo di maturazione del globulo rosso. Il risultato più interessante riporta l'identificazione della proteina annexina VII nella frazione dei neociti, mentre la proteina è assente nelle cellule più mature. L'andamento del profilo di espressione dell'annexina VII suscita particolare interesse perché il suo ruolo nell'eritrocita non è ancora noto ed evidenze sperimentali la associano a processi di vescicolazione. Sono noti in letteratura, infatti, il suo legame alle membrane lipidiche, e la sua presenza, in elevate concentrazioni, nei "raft" lipidici, ovvero agglomerati di proteine di membrana, coinvolti nei processi di endo ed esocitosi. Esperimenti di immunoblotting hanno confermato l'assenza di annexina VII nelle frazioni eritroidi mature di soggetti sani. Per valutare un'eventuale relazione tra l'annexina VII e il rimodellamento cellulare è stato indagato il profilo di espressione della proteina in frazioni cellulari eritroidi giovani e mature di due casi di sferocitosi, rispettivamente caratterizzati da deficit di anchirina e spettina e deficit di banda 3. Infine, essendo la milza l'organo coinvolto nel rimodellamento cellulare del globulo rosso durante l'invecchiamento, le due patologie sono state studiate in pazienti splenectomizzati e non splenectomizzati.

I risultati rivelano un comportamento differente in soggetti affetti da sferocitosi. In particolare, l'annexina VII è presente in tutte le fasi dell'evoluzione cellulare dei globuli rossi affetti da sferocitosi per deficit di banda 3, ed ha un comportamento analogo al controllo sano nei globuli rossi affetti da sferocitosi per deficit di anchirina e spettina. Infine, la milza non incide sulla rimozione dell'annexina VII durante l'invecchiamento cellulare.

I risultati conseguiti fin ora hanno consentito di focalizzare l'attenzione su una proteina, l'annexina VII, la cui funzione non è ancora ben definita, nell'ambito delle sferocitosi. Ulteriori indagini, sia di carattere biochimico che clinico, sono necessarie per studiare il ruolo svolto dall'annexina VII.

### **Caratterizzazione dei target molecolari del fattore trascrizionale TBX1. Sindrome di DiGeorge**

La sindrome di DiGeorge (DGS) è un esempio di disordine nello sviluppo dell'apparato faringeo. Molti casi di DGS sono associati a delezioni cromosomiche eterozigotiche che interessano il cromosoma 22q11.2. Sono tra le più frequenti delezioni cromosomiche associate alla nascita (1 ogni 4000 nati vivi), e si manifestano con un'ampia variabilità fenotipica. Esperimenti condotti su modelli murini mostrano che delezioni eterozigotiche lungo una regione cromosomica di 1 Mb, omologa a quella umana, determinano, nell'animale, anomalie cardiovascolari, caratteristiche della patologia umana. L'unico gene aploinsufficiente nella regione cromosomica studiata è TBX1. Tbx1 appartiene alla famiglia dei fattori trascrizionali che possiedono un dominio T-box di legame al DNA ed è una proteina essenziale allo sviluppo del sistema cardiovascolare sia nella classe tassonomica umana che murina. Inoltre la scoperta di mutazioni puntiformi nel gene TBX1 in pazienti affetti da sindrome di DiGeorge, dimostra il coinvolgimento del gene stesso nella patogenesi DGS. Pur essendo chiaro il coinvolgimento di Tbx1 nell'insorgenza della patologia, non sono ancora noti gli effetti delle mutazioni puntiformi studiate sulla

macchina cellulare. Obiettivo del progetto di ricerca è stata l'identificazione delle proteine la cui espressione dipende dall'attività del fattore trascrizionale Tbx1, con l'ausilio di metodologie di proteomica di espressione differenziale.

Essendo Tbx1 una proteina coinvolta nello sviluppo dei cardiomiociti è stato scelto per lo studio dell'espressione proteica un sistema modello, costituito da cellule indifferenziate embrionali di carcinoma murino (P19CL6) che differenziano a cardiomiociti se indotte con DMSO. La linea cellulare scelta è stata stabilmente trasfettata rispettivamente con un plasmide esprimente Tbx1 in modo costitutivo ed un plasmide vuoto, usato come controllo. Mediante la tecnologia del DIGE (Differential Gel Electrophoresis) sono stati confrontati, da un punto di vista qualitativo e quantitativo i livelli di espressione proteica delle due linee cellulari descritte. In particolare i lisati proteici, purificati da cellule P19CL6 esprimenti Tbx1 e cellule di controllo, sono stati marcati con due fluorocromi, caratterizzati da spettri di emissione ed eccitazione differenti. Le miscele proteiche marcate in modo indipendente sono separate in un unico esperimento di elettroforesi bidimensionale. Le singole immagini relative ai due lisati sono state acquisite a lunghezza d'onda differenti e normalizzate rispetto ad uno standard, costituito da una miscela equimolare dei due estratti confrontati, marcata con un terzo fluorocromo. Affinché l'analisi quantitativa possa avere validità statistica, l'esperimento è stato eseguito utilizzando quattro replicati biologici. L'analisi d'immagine comparativa ha consentito di determinare i cambiamenti nell'espressione di circa 50 proteine, attraverso la misura dell'intensità di fluorescenza di ogni singolo spot. Gli spots di interesse sono stati escissi, sottoposti ad idrolisi *in situ* e le miscele peptidiche ottenute, analizzate per spettrometria di massa tandem. L'identificazione delle proteine presenti negli spots escissi è stata eseguita avvalendosi di opportuni sistemi computazionali.

Lo studio dei prodotti di espressione genica, bersaglio dell'attività trascrizionale di Tbx1 ha condotto all'identificazione di circa 50 proteine, divise in ipo o iper-esprese in conseguenza alla maggiore presenza di Tbx1 nel sistema cellulare. Le proteine trovate sono state classificate in base alla loro attività funzionale: un'ampia parte è coinvolta in processi di nucleotidasi tri-fosfatasi (NTPasi), e di legame a proteine, DNA e RNA. L'attività fosfatasica induce ad ipotizzare un coinvolgimento delle proteine identificate nella regolazione dei processi di trasduzione del segnale. Il legame a DNA e RNA suggerisce un legame con processi di trascrizione e splicing.

Tra le proteine ipoesprese nei lisati proteici di cellule over-esprimenti Tbx1, l'*aldehyde dehydrogenase 2* (Adh2) ha destato particolare interesse, essendo un enzima coinvolto nella ossidazione della retinaldeide ad acido retinoico, la cui concentrazione molare è critica per lo sviluppo dell'apparato faringeo. E' noto che la mancanza di TBX1 nella sindrome DGS determina un aumento della concentrazione dell'acido retinoico. Questo dato avvalorava il risultato scientifico ottenuto. Altra proteina d'interesse è una ciclica coinvolta nella divisione cellulare, *proliferating cell nuclear antigen* (PCNA), ipoespressa nel sistema over-esprimente Tbx1. Il suo trascritto aumenta in seguito a stimolazione con acido retinoico, ma ancora nulla si conosce in merito al comportamento del prodotto proteico.

### **Interattoma della proteina RPS19. Implicazioni nell'anemia di Diamond Blackfan**

L'anemia di Diamond Blackfan (DBA) è una malattia congenita, caratterizzata da una anomala maturazione eritropoietica. Il 25% dei casi patologici per DBA presenta una mutazione puntiforme nella regione cromosomica 19q13.2 che codifica una proteina ribosomiale RPS19. E' stato ipotizzato che le mutazioni trovate incidono sull'espressione proteica e/o determinano una perdita di funzione della proteina stessa, non direttamente connessa alla struttura del ribosoma. RPS19 è, infatti, una proteina strutturale,

componente della subunità ribosomiale 40S. Analogamente ad altre proteine ribosomiali, RPS19 trasloca dal citoplasma al nucleo, dove partecipa alla biogenesi del ribosoma. E' stato dimostrato che in lievito la sua assenza determina una anomala idrolisi dell'RNA con conseguenze sulla biogenesi stessa del ribosoma. Pertanto, RPS19 potrebbe avere, oltre alla sua funzione puramente strutturale, anche un ruolo nella sintesi e maturazione del ribosoma. Sono note, infatti, diverse proteine ribosomiali che svolgono funzioni importanti, avulse dal loro ruolo strutturale: RPL13 e RPL26 sono dei marcatori tradizionali. La funzione biologica di una proteina può essere determinata dalla conoscenza delle interazioni che essa instaura all'interno della cellula. La definizione dell'intorno topologico di una proteina è realizzata mediante approcci sperimentali di proteomica funzionale. Il cDNA del gene RPS19 è stato clonato in un vettore plasmidico, ed espresso come proteina di fusione GST-RPS19. In modo analogo è stata sintetizzata la proteina GST, usata come controllo. Le proteine GST e GST-RPS19 sono state cimentate con estratti totali di cellule umane eritroleucemiche K562. I complessi, risultanti dall'interazione *in vitro* delle proteine di fusione con i propri partner molecolari presenti nell'estratto, sono stati purificati per cromatografia d'affinità, sfruttando l'interazione specifica tra la GST e una molecola di glutathione immobilizzata su una resina. Le componenti proteiche eluite sono state frazionate per SDS-PAGE e visualizzate in seguito a colorazione coomassie. Le bande proteiche presenti nelle due frazioni sono state escisse dal gel e sottoposte ad idrolisi *in situ* con un enzima proteolitico specifico. Le miscele peptidiche ottenute sono state analizzate per spettrometria di massa tandem. L'identificazione delle proteine presenti nelle bande escisse dal gel è stata eseguita avvalendosi di opportuni sistemi computazionali.

Il confronto tra l'insieme di proteine purificate dall'esca GST e dall'esca GST-RPS19 ha permesso di definire l'interattoma della proteina RPS19. Sono state identificate 159 proteine, classificate in base alla loro funzione: ATP/GTPasi, idrolisi/elicas, isomerasi, chinasi, fattori di splicing, costituenti strutturali del ribosoma, fattori trascrizionali, transferasi, trasportatori e specie proteiche leganti DNA/RNA, deidrogenasi, ligasi, peptidasi, fattori di elongazione, recettori, e proteine di funzione non nota. Il lavoro svolto rappresenta un primo approccio globale su larga scala per lo studio dell'interattoma di RPS19. Le 159 specie identificate hanno una localizzazione prevalentemente nucleare e nucleolare, nonostante la purificazione cromatografia sia stata condotta su un estratto proteico totale. Questo sottolinea la funzionalità della proteina di fusione GST-RPS19 anche in presenza del tag. Poiché molte proteine identificate appartengono alla subunità ribosomiale pre-40S e alla 60S è possibile che molte di esse, tra cui RPS19, partecipino all'assemblaggio già nella fase antecedente la maturazione dell'rRNA, nel pre-ribosoma 90S. Quest'ultima è una struttura dinamica che si avvale di 150 proteine non ribosomiali con un ampio intervallo di attività, quali RNA elicasi, GTPasi, nucleasi, ATPasi, chinasi (45, 46). Inoltre sono state trovate, in suddetto studio, 23 delle 31 proteine i cui orologi, in lievito, appartengono alla subunità 90S. Poiché l'interattoma di RPS19 si compone di proteine, quali RPS10, RPS26, RPS23, RPS2 e l'esportina XPO, coinvolte nel trasporto delle subunità ribosomiali dal nucleo al citoplasma, è possibile ipotizzare, anche per RPS19, la medesima funzione. RPS19 mostra, inoltre, possedere molti interattori comuni ad una peptidyl prolin isomerasi, la parvullina, importante nelle prime fasi della maturazione ribosomiale. Infine sono state trovate proteine importanti nel controllo della sintesi proteica, suggerendo anche per RPS19 un ruolo analogo. La definizione dei ligandi molecolari di nota funzione di una proteina può chiarire i processi cellulari in cui la proteina stessa resta coinvolta e far luce sul meccanismo d'azione che la associa alla patogenesi DBA.

## SUMMARY

Proteomics is a modern approach aimed to explain the information contained in genomic sequences in terms of structure, function, and control of biological processes and pathways. Proteomics attempts to study biological process comprehensively by systematic analysis of protein expressed in a cell or a tissue.

Current proteomic investigations are essentially focused on two major areas, expression proteomics, aimed to measure up and down regulation of protein levels, and functional proteomics aimed at the characterisation of multiprotein complexes.

In biomedical application, a comparative approach is usually employed to identify proteins that are up- or down regulated in a disease specific manner for use as diagnostic markers or therapeutic targets. The main analytical steps involve the protein fractionation by mono or bi-dimensional electrophoresis and their identification by mass spectrometer system. The present work has been developed in three main projects:

- ✓ Comparative study of red cell membrane proteins as a function of red cell age
- ✓ Quantitative and differential analysis of the transcription factor TBX1 molecular targets, involved in DiGeorge syndrome
- ✓ Functional analysis of RPS19 interacting proteins.

### **A proteomic study of red cell membrane proteins as a function of red cell age**

Red blood cell functions are made possible by the interaction of the three major cellular components: membrane, haemoglobin and metabolic pathways that allows to modulate oxygen transport, protect haemoglobin from oxidant-induced damage and maintain a stable cell osmotic environment. The red cell membrane consists of a lipid bilayer and integral proteins attached to an underlying protein skeleton. This structure is designed in order to provide the maximum plasticity and deformability so that the erythrocyte can survive in the microcirculation. Defects in horizontal interactions lead to cytoskeletal destabilization. A mild destabilization causes a defective shape recovery, which constitute the molecular defect of pathologies classified under the name of hereditary spherocytosis (HS).

The red blood cell membrane consists of a lipid bilayer studded with integral proteins, supported by an underlying protein skeleton. This structure is designed to provide the maximum plasticity and deformability, in order to optimise the survival of the erythrocyte in the microcirculation. HS can be caused by mutations, which involve five proteins that are implicated in 'vertical' associations between the cytoskeleton and the overlaying lipid bilayer:  $\alpha$  and  $\beta$  spectrin, ankyrin, band 3, and protein 4.2.

The human red cell represents the final differentiation stage of the erythroid series. During its lifespan of about 120 days it demonstrates variation in structure. *In vivo*, older cells are denser than younger. Red blood cell membrane proteins in normal and in HS, as a function of cell age, were characterised using a proteomics.

Red blood cell membrane proteins at different cell densities in healthy controls and HS patients with ankyrin and spectrin or band 3 deficiencies were analysed. The erythrocytes were separated into age-related fractions using self-forming Percoll density gradients. The membrane proteins were separated by 2DE. The first results demonstrate that during cell aging a series of proteolytic products of cytoskeletal proteins is generated such as ankyrin and band 3. The results confirm that lower molecular weight forms occur during red cell maturation. The comparison of red cell membrane proteins at different cell densities showed that annexin VII is present in reticulocytes but it is then lost as the mature cells. The reticulocytes have a greater deformability and capability for vesiculation with the formation of micro vesicular and Nan vesicular 'rafts' than mature red cells. These are

features which are linked to the known functions of annexin VII. Annexin VII is a  $\text{Ca}^{2+}$  and phospholipid binding protein. It belongs to a family of evolutionarily conserved proteins of a bipartite structure with a variable N-terminal domain and a conserved C-terminal one. This latter is responsible for the  $\text{Ca}^{2+}$  and phospholipid binding. In order to understand the influence of remodelling of the red blood cells on the role of annexin VII, several HS cases, characterised by ankyrin and spectrin or band 3 deficiencies, were analysed. Ankyrin and spectrin deficiencies cause a remodelling of the cytoskeleton, while band 3 deficiency does not. In band 3 deficiency, annexin VII is present in the red blood cell membrane during its entire lifespan, while in samples with ankyrin and spectrin deficiency, annexin VII was lost during ageing as observed for healthy controls. In band 3 deficiency vesiculation involves only integral membrane proteins and cytoskeletal ones. This could be the cause of the presence of annexin VII in older cell. Finally, the spleen does not seem to have a relevant role in the removal of annexin VII because the presence of this protein in aged red blood cells show a similar trend in unsplenectomised and splenectomised samples. These studies illustrate how a proteomic approach can be useful in revealing abnormal protein structure-functions relationships in membrane defects of the red blood cells.

### **Proteomic analysis of the transcription factor TBX1 molecular targets. DiGeorge syndrome**

DiGeorge syndrome (DGS) is a classic example of developmental disorder of the pharyngeal apparatus. Most cases of DGS are associated with a heterozygous interstitial chromosomal deletion of the long arm of chr. 22 (del22q11). The deletion is relatively frequent (1:4000 live births) and is associated with considerable phenotypic variability. The first animal model of del22q11 syndrome was developed by engineering a multi-gene chromosomal deletion of the murine homologous region. Later, it was shown that *Tbx1* is the gene haploinsufficient in the multi-gene deletion model. *Tbx1* is a T-box transcription factor required for segmentation of the embryonic pharynx, patterning of the aortic arch, and morphogenesis of the outflow tract of the heart. The finding of TBX1 point mutations in some patients with typical DGS phenotype conclusively demonstrate the role of this gene in the pathogenesis of DGS. Although genetic studies have produced a large amount of knowledge, nevertheless the complexity of the system highlights that several important aspects remain to be elucidated at the level of proteins. Proteomic analysis offers the potential to look at all the major proteins involved in TBX1 mediated pathways and to determine how protein expression differs between different cellular states. In biomedical application, a comparative approach is usually employed to identify proteins that are up- or down regulated in a disease specific manner for use as diagnostic markers or therapeutic targets. In these studies, a reliable analysis of quantitative changes in protein expression is crucial. Identification of proteins whose expression is modified by *Tbx1* will be achieved by performing a comparative two dimensional SDS-PAGE gel (2DE SDS-PAGE) analysis of the lysates originated from *Tbx1*-free cell lines vs. *Tbx1*-expressing cell lines. An efficient comparative procedure based on dual fluorescent labelling will be performed on TBX1 stably transfected P19CL6 (mouse embryonal carcinoma) cells that represent a tissue culture model for cardiomyocyte differentiation (P19-TBX1-TEV-PA). These cells differentiate *in vitro* into beating cardiomyocytes in the presence of DMSO. Because of *Tbx1* is required in cardiomyocyte precursors, these cells may be a convenient model to study *Tbx1* transcriptional targets. According to the DIGE procedure the two protein lysate samples were labelled *in vitro* using two different fluorescent cyanine minimal dyes (Cy3 and Cy5, respectively) differing in their excitation and emission wavelengths, then mixed before isoelectrofocusing (IEF) and separated on a single 2DE SDS-PAGE. After consecutive excitation of both wavelengths, the images

were overlaid and subtracted (normalised), whereby only differences (e.g. up- or down regulated proteins) between the two samples were visualised. A third cyanine dye (Cy2) was used to label a mixture of both samples in order to include them in the 2DE SDS-PAGE analysis as internal standard. By performing an accurate image analysis it was possible to visualise significant differences in some of the protein spots present on the gel. About 50 proteins resulted to be quantitatively increased or decreased in the P19-Tbx1-TEV-PA cell line, compared to the control cell line. Protein spots were excised from the preparative gel, digested by proteolytic enzyme and identified by mass spectrometer procedure. The identified proteins were classified and grouped on basis of specific function. Among these was identified a group involved in cellular signalling control, such as NTPase, and transcription regulators, such as DNA and RNA binding proteins. Among the down-regulated proteins two species are directly related to Tbx1 function: *aldehyde dehydrogenase 2* (Adh2) and *proliferative cell nuclear antigen* (PCNA). Adh2 catalyzes the oxidation of retinaldehyde into retinoic acid (RA). This later result accords to other literature data, where the down-regulation of Tbx1 has been reported associated with an increasing of RA intracellular levels. Even PCNA seems too related to RA metabolism: indeed Martin et al. verified that the transcription of PCNA gene is stimulated by RA.

### **Analysis of RPS19 interacting proteins. Diamond Blackfan anaemia**

Diamond Blackfan Anaemia (DBA) is a congenital disease characterized by defective erythroid progenitor maturation. Mutation on chromosome 19q13.2 encoding ribosomal protein (RP) S19 account for 25% of patients.

Deregulation of RPS19 has thus been surmised as the cause of this defective erythropoiesis, though the link between them is still unclear. Two hypotheses have been proposed: altered protein synthesis, and loss of unknown functions not directly connected with RPS19's structural role in the ribosome. RPS19 is a structural component of the ribosomal 40S subunit. Like other ribosomal proteins (RPs), RPS19 translocates from the cytoplasm to the nucleus where it participates in ribosome biogenesis. In yeast its absence is associated with abnormal rRNA cleavage and defective ribosomal biogenesis. It thus appears that RPS19, in addition to its structural role in the ribosome, is involved in ribosome biogenesis, specifically in rRNA processing, and possibly, in translation. These functions are probably assisted by interaction with different protein substrates.

Functional proteomic procedures were used to look for proteins interacting with RPS19 and thus secure additional information regarding its function and regulation.

The human RPS19 cDNA was amplified by RT-PCR and cloned into pGEX-4T-1 to generate plasmid pGEX-RPS19. GST and GST-RPS19 fusion proteins were expressed in *E. Coli* cells. K562 cell extracts were pre-incubated with GST and GST-RPS19 resin. The proteins eluted from the GST-RPS19 or the negative control GST resins were fractionated on a 13 cm 8-18% SDS-PAGE and revealed by colloidal Coomassie stain. SDS-PAGE indicated that the RPS19-associated complexes contained hundreds of proteins spanning a broad molecular weight range. The GST-RPS19 and GST gel lanes were totally cut to create sixty-five 2 mm slices per lane. Peptide mixture samples from each lane were analysed by mass spectrometry. Mass spectrometry data were then analysed with the Mascot software on the NCBI human protein sequence database. To select proteins that interact specifically with RPS19, we subtracted species common to the GST and GST-RPS19 lanes.

The 159 human proteins identified in this study were divided into Gene Ontology functional groups as: NTPases (ATP- and GTP-ases), hydrolase/helicases, isomerases, kinases, splicing factors, structural constituents of ribosome, transcription factors, transferases, transporters, DNA/RNA-binding protein species, dehydrogenase, ligase, peptidase, protein

receptor, elongation factor and protein of still unknown function. They were also grouped according to their Gene Ontology cellular localization and the biological processes in which are involved.

This study represents the first global, high-throughput functional proteomic approach to identify the proteins that interact with RPS19. Interestingly, most proteins reported in this study, such as nucleolar or ribosomal proteins, play a role in processes related to RPS19. It should be stressed that we used a total cell lysate and not a nuclear extract and that the complex formation was extracellular. Proteins abundant in cytoplasm were not found. This suggests that the structure of the recombinant RPS19 protein is functionally suitable to recruit multiple cellular partners. The identification of RPs belonging to the small (40S) and the large (60S) subunit suggests that we have purified components of the pre-ribosome (90S), the structure formed before processing of the pre-rRNA. The pre-ribosome is a highly dynamic structure that comprises more than 150 non-ribosomal proteins with various activities, including nucleases, RNA helicases, GTPases, ATPases, kinases, etc. The interaction with most of the RPs essential for the transport of the small subunit from the nucleus to the cytoplasm (i.e. RPS10, RPS26, RPS3, and RPS2) and to the exportin XPO, known to control the 40S and 60S export, suggests a role for RPS19 in this process. Many interactors are shared between RPS19 and parvulin, a peptidyl prolyl isomerase involved in early ribosome biogenesis. A further group of interactors includes proteins controlling protein synthesis, such as proteins involved in cotranslational translocation (such as signal recognition particle srp68) and translation regulators, such as IGF2BP1 and STAU1. Other ribosomal proteins (i.e. RPS13 and RPL26) are known to regulate translation of specific transcripts. It is intriguing that RPS19 could have a similar role.

# Capitolo 1

## INTRODUZIONE

### 1.1 Premesse scientifiche

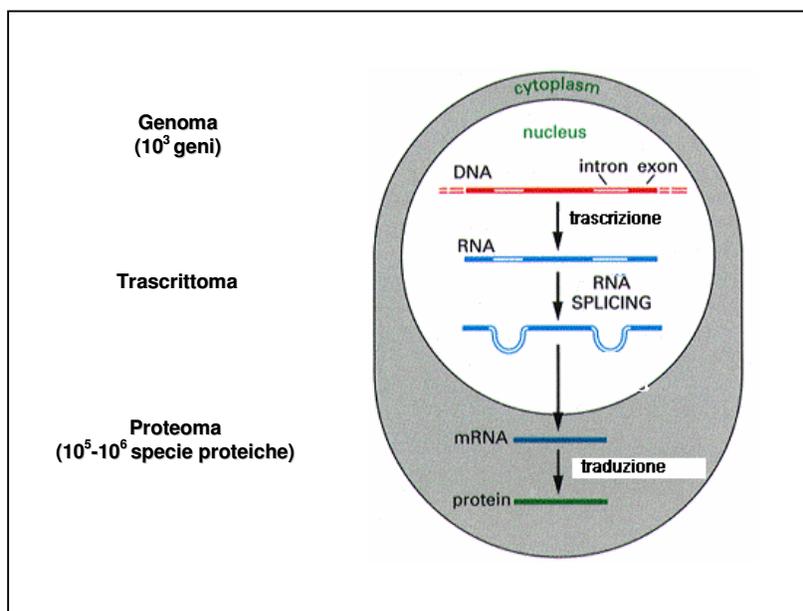
La proteomica è l'insieme delle metodologie sperimentali necessarie a tradurre le informazioni contenute in una sequenza genomica, in termini di struttura, funzione e controllo di processi biologici, attraverso l'analisi sistematica di tutte le proteine espresse non solo in una cellula intera o in un tessuto, ma anche in fluidi biologici. Si propone, dunque, lo studio della totalità delle proteine nella loro complessa attualità funzionale, al fine di comprendere il funzionamento della macchina cellulare, sviluppare nuove strategie di diagnosi precoce, identificare bersagli terapeutici, con conseguente implemento, sia nella ricerca di farmaci innovativi sia nelle successive applicazioni, volte a valutarne gli effetti terapeutici o tossici [1]. Gli studi genetici hanno prodotto innumerevoli informazioni in merito a molte patologie. Tuttavia molto resta ancora da chiarire circa le vie attraverso cui un gene, o un insieme di geni, determinano l'insorgenza di una patologia. Uno studio di tipo proteomico può complementare le informazioni ricavate da una analisi genetica, superando i limiti di altri approcci e fornendo informazioni circa interazioni, funzioni e fluttuazioni di un corredo proteico, elementi che hanno un ruolo cruciale nel sistema cellulare. Rispetto ad un classico studio di tipo chimico o biochimico, la proteomica si propone la descrizione dei complessi sistemi proteici, dinamici ed interattivi, piuttosto che la caratterizzazione di proteine selezionate, antepoendo un punto di vista globale, ad uno specifico [2-3].

Studiare il corredo proteico nella sua globalità vuol dire considerare le modifiche, in termini quantitativi e qualitativi, che possono essere apportate da fattori endogeni, quali modulazione della trascrizione, splicing alternativi, modificazioni post traduzionali, interazioni con altre proteine o acidi nucleici. Il livello di espressione di un gene, di traduzione di un trascritto, e il conseguente numero di prodotti proteici per gene o trascritto, può variare in un intervallo molto ampio di ordini di grandezza in dipendenza dalle condizioni cellulari e dall'ambiente esterno. Fenomeni di splicing determinano una corrispondenza non biunivoca tra gene e proteina. Infine una sequenza polipeptidica può subire diversi tipi di modificazione post traduzionale, tra cui fosforilazioni, acetilazioni, glicosilazioni, ubiquitinazioni [4].

Tutti questi fattori non sempre correlati in modo diretto con la fisiologia e la patologia, modificano aspetti peculiari, quali ad esempio l'emivita molecolare, che non sono prevedibili dalla struttura dell'acido nucleico o del messaggero da cui derivano.

Si conclude, che un proteoma ha un maggiore ordine di complessità rispetto ad un genoma, ed oppone alla natura di quest'ultimo essenzialmente statica una forte variabilità (figura 1). Catturare e descrivere questo stato dinamico rappresenta una sfida per lo sviluppo delle nuove tecnologie.

Un'importante distinzione riguarda l'obiettivo specifico che un'indagine proteomica intende perseguire. Gli scopi dell'analisi possono essere molteplici: misura della variazione di espressione di una specie proteica e/o determinazione della sua localizzazione e/o interazioni. Gli attuali studi proteomici, pertanto, sono focalizzati su due principali aree: la proteomica di espressione per la definizione qualitativa e quantitativa dell'aumento e/o diminuzione dei livelli di espressione proteica; la proteomica funzionale volta a stabilire le interazioni, che le proteine stabiliscono nel proprio intorno cellulare [5].



**Figura 1.** Un proteoma ha un maggiore ordine di complessità rispetto ad un genoma.

## 1.2 Proteomica funzionale

Definire la funzione di una proteina, oggetto di studio, significa associare ad essa, una posizione definita all'interno di una cascata di eventi cellulari.

E' ben noto, infatti, che molte delle funzioni cellulari non sono svolte dalle proteine come entità individuali, ma in combinazione con altre componenti proteiche che, stabilendo delle interazioni rapide e specifiche intervengono e condizionano la realtà funzionale di una data proteina [1]. Oltre alla formazione di complessi multiproteici stabili, le proteine si associano in modo transiente ai loro substrati per modificarli, esercitare un effetto regolativo anche solo per ragioni di tipo sterico, o per traslocarli in un differente compartimento cellulare. L'iterazione di questi processi realizza delle vere e proprie reti molecolari, dotate di nodi e diramazioni che, in definitiva, creano circuiti in cui scorre la trasmissione delle informazioni. La comprensione delle funzioni di una proteina, così come la definizione dei meccanismi molecolari all'interno di una cellula dipende quindi dall'identificazione dei partners con cui interagisce. L'associazione di una proteina sconosciuta con partners che appartengono ad uno specifico complesso multiproteico, coinvolto in un determinato meccanismo cellulare potrebbe, infatti, essere fortemente indicativa della sua funzione biologica [6]. Inoltre la definizione delle interazioni tra proteine all'interno della cellula potrebbe essere di grande aiuto anche nella descrizione dettagliata dei percorsi di signalling cellulare [7].

Le proteine che si organizzano in questi sistemi multiproteici costituiscono dei sub-proteomi che non possono essere rivelati con il metodo classico del "doppio ibrido". La tecnica del doppio ibrido consiste in una selezione genetica che rivela interazioni di tipo binario, attraverso l'uso di una proteina-esca che, formando un complesso con una proteina-preda, attiva la sintesi di una proteina-segnales. La strategia proteomica si avvale della possibilità di esprimere, in un sistema cellulare, la proteina d'interesse, che definisce esca, in fusione ad una sequenza polipeptidica, un *tag*, quale ad es. un epitoco riconosciuto da anticorpo o un enzima che lega uno specifico substrato. I complessi, che l'esca forma con i suoi ligandi all'interno di un estratto proteico, possono essere purificati

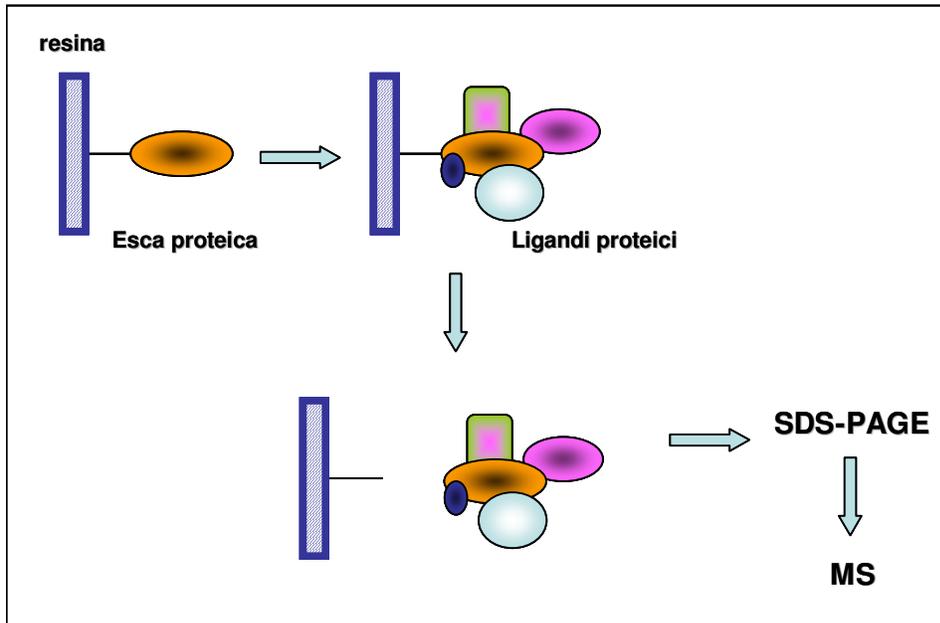
per cromatografia d'affinità sfruttando il legame specifico del tag proteico (figura 2) [8]. Con un sistema di espressione procariotico, la proteina esca può essere espressa come proteina ibrida ricombinante, in fusione con l'enzima *Glutathione S-Transferase* (GST), oppure può essere modificata con una coda di istidine (*poli-His*), biotinilata, etc.

Purificata dal lisato procariotico in cui è stata espressa e cimentata con opportuni estratti proteici, la proteina ricombinante instaura, *in vitro*, le interazioni con i suoi ligandi. I risultanti complessi multiproteici sono purificati per cromatografia d'affinità che utilizza una resina di sefarosio, derivatizzata con molecole di glutathione per legare le proteine ibride con GST, o con ioni Nichel per le proteine modificate con code di istidine. L'uso di tutti questi marcatori di affinità è applicabile ad un grande numero di proteine poichè si osserva solo un minimo e trascurabile effetto sulla struttura terziaria (e quindi sull'attività della proteina esca) che potrebbe ostacolare la stabilità del complesso. Il successo dell'approccio basato su tecniche di affinità dipende tuttavia dall'assenza di eccessive interazioni non specifiche, correlata alla selettività del riconoscimento *in vitro* fra l'esca ed il partner. E' necessario operare dei passaggi di pre-purificazione al fine di minimizzare le contaminazioni di fondo. Strategie alternative fondano su tecniche di immunoprecipitazione. La proteina esca è espressa in fusione con un tag costituito da un epitopo peptidico per il quale è disponibile un buon anticorpo (FLAG, HA, Myc, etc.). L'esca, trasfettata nella linea cellulare d'interesse, formerà *in vivo* il complesso multiproteico interagendo con i suoi partners specifici. Il complesso è immunoprecipitato utilizzando un anticorpo specifico per l'epitopo legato alla proteina esca. Sebbene l'approccio dell'immunoprecipitazione sembri produrre dati significativi, la reattività incrociata con antigeni aspecifici, così come il legame non specifico delle proteine, dei tag o del supporto insolubile agli anticorpi possono portare a falsi positivi.

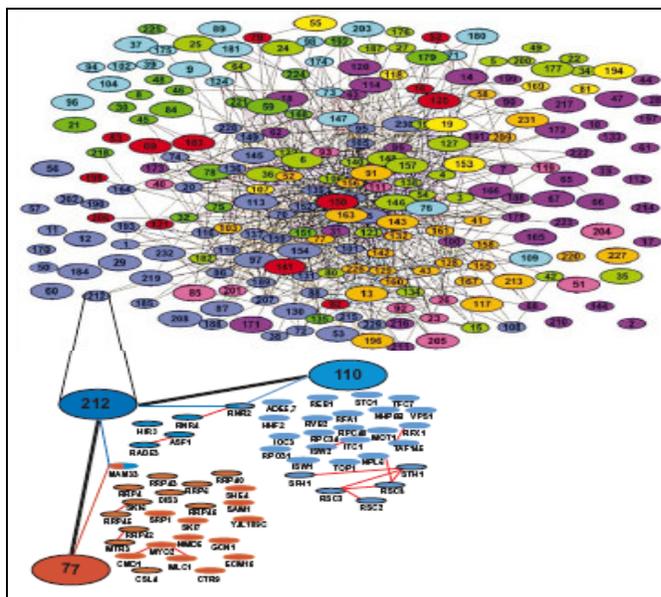
Un sistema innovativo propone l'impiego di una duplice purificazione mediante cromatografia affinità (Tandem Affinity Purification tag system, TAP-tag) per lo studio di complessi multiproteici in alte rese e in condizioni native. Secondo questa metodologia la proteina esca è modificata con due *tags* differenti (la Proteina A, riconosciuta dalle Immunoglobuline G (IgG) e un peptide capace di interagire fortemente con la Calmodulina), separati da una sequenza idrolizzata in presenza di un enzima specifico: un sito di idrolisi per la proteasi del virus del tabacco (TEV). Il costrutto, introdotto transientemente o stabilmente nelle cellule o negli organismi di elezione, permette di isolare complessi multiproteici ai loro livelli naturali.

Le proteine costituenti il complesso purificato con una delle tecniche descritte sono quindi frazionate mediante SDS-PAGE ed identificate utilizzando metodologie di spettrometria di massa.

Con le tecniche di purificazione elencate è stata descritta una organizzazione di 1784 complessi proteici in lievito, ciascuno dei quali condivide con altri alcuni dei propri componenti molecolari. E' stata realizzata una mappa topografica delle interazioni, che ha evidenziato funzioni, finora sconosciute, per molte proteine note (figura 3) [6].



**Figura 2.** La strategia utilizzata in proteomica funzionale prevede la derivatizzazione di una resina con una proteina esca fusa ad un tag, la formazione e l'eluizione del complesso mediante cromatografia d'affinità. Il complesso è quindi frazionato mediante elettroforesi e le sue componenti identificate per spettrometria di massa.



**Figura 3.** Complessa rete di interazioni proteiche definita mediante metodologie di proteomica funzionale [6].

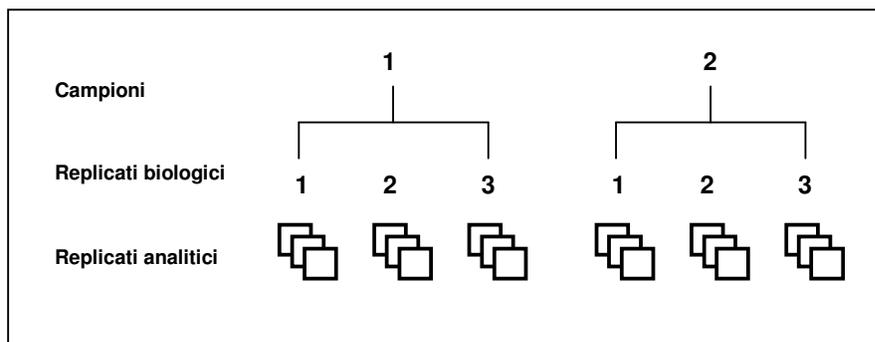
### 1.3 Proteomica differenziale

Consentendo l'analisi simultanea di migliaia di proteine, tecnologie proteomiche sono utilizzate per studi di tipo comparativo. E' possibile paragonare, da un punto di vista sia qualitativo che quantitativo, i profili di espressione dei proteomi di una stessa cellula o di uno stesso organismo, in diverse condizioni. Nelle applicazioni biomediche un'analisi di tipo comparativo è utilizzata per identificare quelle proteine, i cui livelli di espressione si mostrano alterati in un fenotipo patologico al fine di definire nuovi marcatori diagnostici o bersagli terapeutici. La tecnica analitica d'elezione per studi di tipo comparativo è l'elettroforesi bidimensionale (2DE), che separa le proteine di miscele complesse mediante due proprietà intrinseche, quali punto isoelettrico e peso molecolare. Sfruttando due passaggi analitici successivi ed ortogonali (isoelettrofocalizzazione, IEF, e SDS PAGE) è possibile assegnare ad ogni molecola proteica una posizione all'interno di una mappa definita da due coordinate: pI e peso molecolare [9]. Sebbene la tecnica sia stata introdotta già negli anni settanta, il suo impiego effettivo è tuttavia recente. Condizione necessaria al suo impiego in proteomica differenziale è la riproducibilità delle mappe bidimensionali. Questa condizione è stata soddisfatta solo con l'introduzione delle "immobiline" (IPGs *immobilized pH gradients*) per la formazione del gradiente di pH lungo il gel d'acrilammide [10]. Una convenzionale isoelettrofocalizzazione era condotta in presenza di composti anfoteri che si distribuivano lungo il gel d'acrilammide in funzione del loro pI, sotto l'effetto di un campo elettrico. Le immobiline sono invece composti non anfoteri con valori di pK differenti legati covalentemente ai monomeri d'acrilammide. Pertanto l'applicazione del campo elettrico determina la migrazione esclusiva del campione in esame. L'impiego delle immobiline in 2DE consente inoltre la possibilità di frazionare una maggiore quantità di campione senza perdite in risoluzione. Un ulteriore miglioramento del potere risolutivo legato tecnica risiede nell'introduzione di gradienti su intervalli di pH molto stretti (IPGs *narrow*).

L'analisi differenziale prevede la visualizzazione delle specie proteiche lungo le mappe bidimensionali. La scelta del metodo di visualizzazione è cruciale per l'analisi quantitativo-comparativa. I fattori che determinano suddetta scelta risiedono nella sensibilità e nella linearità dell'intervallo di rilevabilità. La maggiore sensibilità, intesa come la minima quantità di proteina detectata, rende la colorazione con AgNO<sub>3</sub> la tecnica d'elezione per la visualizzazione di proteine. Tuttavia solo una procedura di colorazione, tra quelle convenzionali, che prevede l'impiego del Coomassie Colloidale presenta un intervallo di rilevabilità lineare, inteso come lo spettro di concentrazioni in cui il colorante riesce ad apprezzare differenze quantitative, fornendo un segnale proporzionale all'abbondanza proteica. Una procedura di colorazione con AgNO<sub>3</sub> presenta un limite inferiore di rilevabilità tra 0.1 e 1 ng ed un intervallo di linearità di risposta di un ordine di grandezza. Una procedura di colorazione con Coomassie Colloidale rivela quantità proteiche non inferiori ad 1µg con un intervallo di linearità di risposta da 1 a 3 ordini di grandezza. I protocolli di colorazione di nuova generazione prevedono l'utilizzo di coloranti fluorescenti, quali ad esempio il Sypro Ruby che superano i limiti dei protocolli convenzionali, sia in termini di sensibilità che in linearità di risposta [11-12].

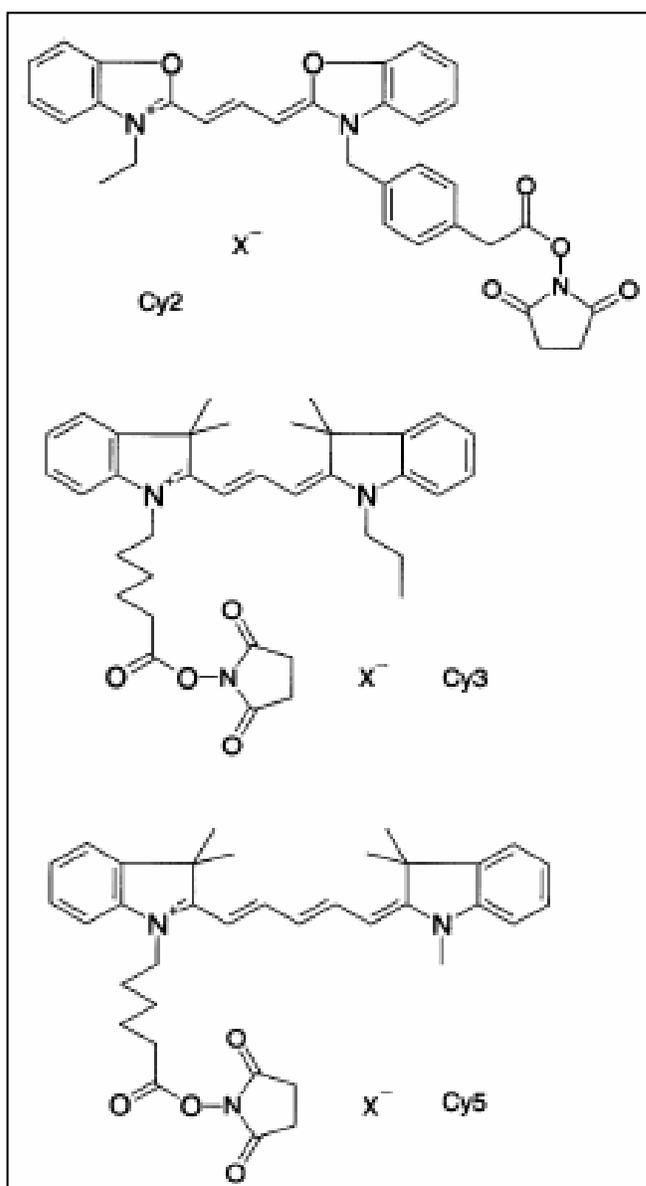
Le mappe ottenute sono acquisite da un densitometro: ad ogni pixel è associato un valore di assorbanza che è proporzionale alla concentrazione della proteina presente nel pixel. Sono stati introdotti numerosi software che assegnano gli spots proteici presenti sul gel, sommano i valori di ogni pixel contenuto nello spot risalendo così ad un volume per ogni spot. Detto volume è diviso per la somma dei volumi di tutti gli spots rivelati e il risultato è un valore relativo espresso come percentuale dell'intero volume rilevato sul gel [13].

La riproducibilità di tali mappe rappresenta un punto chiave per il loro utilizzo in indagini di tipo comparativo. I risultati di un'indagine differenziale richiedono una validazione di tipo statistico. Infatti, sono molte le fonti di variabilità che possono incidere sulle differenze, evidenti nelle mappe bidimensionali prodotte: variazioni analitiche dovute al trattamento del campione, ai procedimenti di colorazione o acquisizione dell'immagine o variazioni biologiche dovute all'ambiente in cui il campione è stato prodotto, processato e conservato. Un disegno sperimentale valido annulla l'impatto che le diverse fonti variazionali analitiche e biologiche hanno sul risultato sperimentale e le minimizza con un'accurata analisi statistica dei dati prodotti. Il disegno sperimentale per studi differenziali ha una struttura gerarchica (figura 4): per confrontare due campioni (stato patologico contro quello fisiologico di una cellula) è importante analizzare da tre a nove replicati biologici per ogni campione, ed eseguire da tre a nove replicati analitici per ogni replicato biologico. I replicati biologici annullano la variabilità biologica, i replicati analitici quella sperimentale [14]. In questo modo le mappe possono essere comparate per rilevare differenze statisticamente significative nell'espressione proteica. L'incremento del numero di replicati valida l'indagine, ma ne aumenta la difficoltà d'analisi [15].



**Figura 4.** Struttura gerarchica per studi differenziali: è necessario eseguire da tre a nove replicati biologici per ogni campione, e da tre a nove replicati analitici per ogni replicato biologico.

L'introduzione di una tecnica innovativa 2D DIGE (Differential Gel Electrophoresis) conferisce a studi di proteomica comparativa una maggiore accuratezza [16]. La tecnica prevede la formazione di un legame covalente tra proteine e coloranti fluorescenti: Cy2, Cy3 e Cy5. I fluorofori hanno analoga struttura chimica e possiedono un gruppo N-idrossi-succinidil-estere in grado di formare un legame ammidico con l' $\epsilon$ -ammino gruppo della catena laterale delle lisine. Un atomo di azoto quaternario presente nella loro struttura reintroduce la carica positiva della lisina (figura 5). I campioni proteici da confrontare sono marcati con due dei tre fluorofori e frazionati nel medesimo esperimento di elettroforesi bidimensionale. I fluorofori hanno, infatti, spettri d'emissione e d'assorbimento non sovrapponibili, il che consente di acquisire due immagini rispettivamente per i due campioni da raffrontare, avendo tuttavia eseguito una sola corsa elettroforetica. Ne consegue un più semplice e rapido protocollo sperimentale rispetto a quello classico. La co-migrazione dei due campioni e la colorazione covalente antecedente la separazione minimizzano le variazioni sperimentali e annullano la necessità di eseguire replicati analitici. La presenza di uno standard interno, invece, riduce il numero di replicati biologici necessari all'indagine statistica.

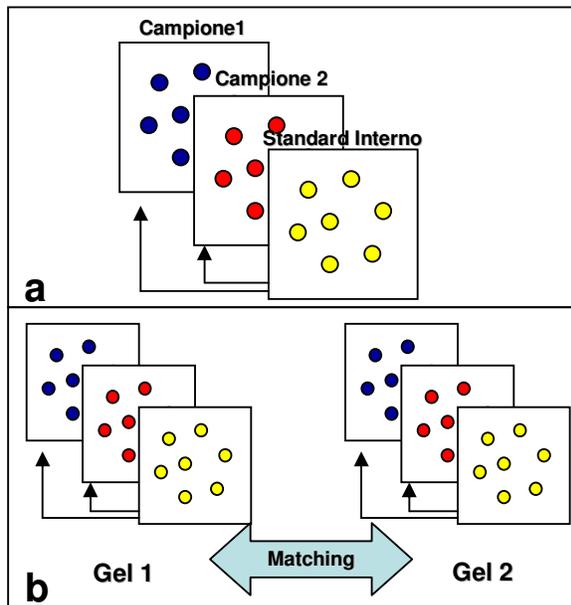


**Figura 5.** I fluorofori Cy2, Cy3, Cy5 hanno analoga struttura chimica ed un gruppo N-idrossi-succimidil-estere in grado di formare un legame ammidico con l' $\epsilon$ -ammino gruppo della catena laterale delle lisine.

Lo standard interno consta di quantità proteiche uguali dei replicati biologici del campione. Le proteine dello standard interno sono legate ad un terzo fluoroforo. Il segnale di fluorescenza relativo ad ogni spot proteico è normalizzato rispetto a quello relativo allo stesso spot dello standard interno. L'analisi quantitativa è statisticamente accurata.

La procedura descritta, nota come *minimal labeling*, consente la marcatura del 3% delle proteine totali di un estratto. Queste ultime, legando covalentemente il fluoroforo, saranno caratterizzate da un peso molecolare superiore di 500Da rispetto alle proteine della stessa specie non marcate. L'analisi d'immagine (figura 6) necessaria per confrontare i campioni in esame prevede due fasi consecutive:

- ✓ co-detection intra gel (6a)
- ✓ matching inter-gel (6b).



**Figura 6.** Il disegno sperimentale per studi differenziali che utilizzano la tecnologia DIGE prevede due fasi consecutive *co-detection intra gel* (a) e *matching inter-gel* (b).

Nella prima fase le tre immagini acquisite in intervalli spettrali differenti sono sovrapposte. Ogni spot è intrinsecamente legato allo stesso spot presente nello standard interno. Infatti, il volume di uno spot inteso come la somma dei valori dei pixel in uno spot sottratto del rumore di fondo è confrontato direttamente con il suo standard interno. L'abbondanza di uno spot in un campione è espressa come rapporto rispetto allo standard interno (normalizzazione).

Nella fase successiva, poiché tutti replicati contengono lo stesso standard interno, è possibile determinare il rapporto dell'abbondanza proteica tra i due campioni. La presenza dello standard interno annulla l'impatto che variazioni sperimentali hanno sull'esperimento. I rapporti campione/standard interno producono un plot di abbondanza relativa standardizzata per ogni specie proteica lungo il gruppo di replicati. In questo modo il confronto tra i due campioni è più accurato. Un test statistico (T-Student) valida il dato, assegnando un intervallo di confidenza  $P < 0,01$  (dove  $P$  è la probabilità che l'evento sia casuale) all'interno del quale il rapporto di abbondanze relative, espresso come il rapporto tra i volumi dei segnali di fluorescenza registrati ( $R = V_1/V_2$ ) abbia un significato biologico ( $R > 1,3$  o  $R < 1,3$ ).

## 1.4 Spettrometria di massa e proteomica

### 1.4.1 Strumentazione

La spettrometria di massa si impone come tecnica elettiva per le identificazioni proteiche nel panorama scientifico proteomico. Ciò è dovuto allo sviluppo di strumentazioni innovative e all'accumulo di una notevole quantità di informazioni in merito a sequenze geniche e proteiche, raccolte in banche dati. La spettrometria di massa si avvale di numerose e differenti strumentazioni, caratterizzate da vantaggi e limiti. Gli spettrometri di massa si compongono di sorgenti, necessarie alla ionizzazione e volatilizzazione del campione in esame, ed un analizzatore che misura il rapporto massa/carica dell'analita, ed

un rivelatore che determina il numero di ioni per ogni rapporto massa/carica trovato. Se in passato il suo impiego era limitato all'analisi di molecole di piccole e medie dimensioni, oggi è una tecnica fondamentale per lo studio delle macromolecole biologiche. Il limite, dovuto ai sistemi di analisi, risiedeva nella difficoltà di trasferire biomolecole polari cariche, di alto peso molecolare, allo stato gassoso. Nuove tecniche di ionizzazione hanno aperto la strada allo studio delle macromolecole biologiche mediante spettrometria di massa: *electrospray ionization* (ESI) e *matrix-assisted laser ionization* (MALDI) [17-19].

Ambedue le tecniche di ionizzazione sono accoppiate ai seguenti analizzatori: *time of flight* (TOF), *quadrupole* (Q), *ion trap* (IT). Le configurazioni comunemente selezionate in base alle esigenze sperimentali sono: MALDI TOF, ESI IT, ESI Q, ESI Q-TOF.

Nella tecnologia MALDI il campione è analizzato in fase solida. La tecnica di ionizzazione, inducendo la formazione preferenziale di ioni monocarica, consente l'analisi diretta di miscele complesse. Il campione in esame co-cristallizza in presenza di una matrice, un composto aromatico acido, che media il trasferimento di energia da un raggio laser (UV o raramente IR) al campione stesso. Il sistema aromatico della matrice è eccitato dal laser e, tornando allo stato fondamentale, trasferisce energia al campione con essa co-cristallizzato. Detta energia è sufficiente a promuovere la formazione di uno ione monocarica  $MH^+$  e ad indurre il desorbimento dal cristallo originario. A definire una tecnica di spettrometria di massa non è solo il metodo di ionizzazione scelto, ma anche il tipo di analizzatore, in cui gli ioni prodotti sono separati in base al rapporto massa/carica. Ad una sorgente MALDI è accoppiato un analizzatore a tempo di volo (TOF), fisicamente costituito da un tubo d'acciaio, di lunghezza compresa tra 1 e 2 metri. Il principio su cui si realizza la separazione sfrutta il fatto che gli ioni generati in sorgente, avendo energia cinetica  $1/2mv^2$ , si muovono con velocità inversamente proporzionali alla radice quadrata della massa e pertanto raggiungeranno il rivelatore in tempi successivi, in ordine di massa crescente.

Nella tecnologia ESI, il campione è analizzato in fase liquida. Un flusso di gas inerte in una camera di ionizzazione, ad alta temperatura e a pressione inferiore a quella atmosferica, favorisce l'evaporazione del solvente, che contiene l'analita carico. L'evaporazione del solvente porta ad una drastica diminuzione del volume: le forze repulsive tra molecole cariche prevalgono rispetto alle forze di tensione superficiale, determinando una repentina nebulizzazione. Un'opportuna differenza di potenziale indirizza gli ioni verso l'analizzatore. Durante il processo di ionizzazione, le molecole proteiche e/o i peptidi vengono protonati da un numero variabile di protoni, in corrispondenza di residui basici o all'estremità N-terminale. Si forma una popolazione di residui a carica multipla  $(M+nH)^{n+}$  che, separati da un analizzatore, danno luogo a distribuzioni di tipo gaussiano dei segnali  $m/z$ , ciascuna delle quali può essere trasformata su scala di massa reale da un algoritmo matematico.

Poiché l'analita è in fase liquida, la tecnica di ionizzazione descritta è accoppiata con una di frazionamento a monte: la cromatografia liquida. Questo tipo di strumentazione è generalmente configurato con un analizzatore quadrupolare o con una trappola ionica. Il primo consta di quattro barre metalliche con geometria cilindrica, alle quali è applicato un campo elettrico costante. Il movimento dello ione dipende da un campo elettrico oscillante a radio-frequenze. Solo lo ione con un definito rapporto  $m/z$  percorrerà, ad un dato valore del campo elettrico, una traiettoria stabile che gli consente l'uscita dal quadrupolo. Nella trappola ionica l'idea di base è opposta. E' un quadrupolo chiuso, al cui interno gli ioni si muovono con una traiettoria stabile, per effetto di un campo elettrico. Gli ioni con un dato rapporto  $m/z$  sono selezionati dal campo elettrico che, rendendo instabile la loro traiettoria all'interno dell'analizzatore li spinge ad uscire [20-21].

## 1.4.2 Strategie

La strategia del *peptide mass fingerprinting* trova larga applicazione in tutte le problematiche correlate all'identificazione di proteine da miscele complesse, separate solitamente su gel SDS-PAGE mono e bidimensionali. Il principio fondamentale su cui esse si basano sfrutta l'idea che il pattern proteolitico di ciascuna proteina sia univocamente dettato dalla specifica sequenza amminoacidica, e che pertanto risulti identificativo della proteina in esame [22]. La procedura prevede l'idrolisi chimica o enzimatica direttamente sulle bande proteiche escisse da gel e la successiva analisi di massa dei peptidi rilasciati, mediante MALDI MS. I pesi molecolari ottenuti sono introdotti in opportuni programmi [23], sviluppati per l'identificazione di proteine, attraverso una ricerca in banche dati telematiche. Qualora la sequenza della molecola in esame sia presente nella banca dati consultata, e la determinazione dei pesi molecolari dei peptidi rilasciati sia stata realizzata con sufficiente accuratezza, l'identificazione della proteina risulta immediata ed univoca. La risposta può essere ambigua e non univoca se una serie di parametri non sono soddisfatti: numero di peptici non sufficiente, presenza di contaminanti peptidici dovuti alla manipolazione del campione, scarsa risoluzione elettroforetica. In tal caso sono utilizzate tecniche di spettrometria di massa tandem accoppiate alla cromatografia liquida. Le miscele peptidiche prodotte dall'idrolisi chimica o enzimatica sono separate mediante HPLC capillare e le frazioni eluite dalla colonna sono direttamente introdotte nella sorgente ESI dello spettrometro. Gli ioni isolati da un primo analizzatore passano in una camera di collisione. Gli ioni frammento prodotti sono selezionati da un secondo analizzatore. I frammenti carichi prodotti danno informazioni in merito alla struttura primaria del peptide selezionato. La strumentazione scelta per questo tipo di indagine prevede l'utilizzo di un accoppiamento tra due analizzatori separati da una camera di collisione oppure di una trappola ionica. In tal caso le banche dati sono interrogate con dati relativi non solo alle masse peptiche, ma anche alla loro sequenza. E' stato dimostrato che due soli peptidi sono sufficienti per identificare una proteina in modo non ambiguo [24].

## 1.5 Scopo del progetto di tesi

Il lavoro svolto durante il progetto di tesi è stato, dunque, incentrato sullo sviluppo di metodologie innovative in proteomica clinica ed articolato come segue:

- ✓ Studio delle proteine di membrana di RBC mediante l'impiego di metodologie di proteomica differenziale.  
L'organizzazione strutturale e la composizione biochimica sono determinanti per la funzionalità della membrana eritroide. Da esse dipendono, infatti, parametri cellulari, quali la forma, il volume cellulare medio, la resistenza osmotica e la deformabilità della cellula stessa. Difetti nelle interazioni proteina-proteina lungo la membrana, o tra membrana e citoscheletro, o nella composizione proteica di membrana, determinano anomalie nei rapporti superficie-volume e nelle concentrazioni dei cationi monovalenti, destabilizzando il doppio strato lipidico con conseguente perdita del materiale proteico. La cellula, pertanto, assume una forma sferica che modifica proprietà della membrana, quali plasticità e deformabilità, indispensabili alla sua sopravvivenza nel microcircolo e alla sua resistenza osmotica. Le anomalie elencate costituiscono le basi molecolari di un'ampia classe di patologie, note come sferocitosi. Per poter studiare l'organizzazione strutturale, la composizione e l'evoluzione temporale della membrana eritroide sono stati confrontati i proteomi di due popolazioni cellulari: cellule eritroidi immature ed adulte. Queste ultime, infatti, pur derivando dal medesimo campione, differiscono per la diversa densità, deformabilità e comportamento in processi cellulari.
- ✓ Identificazione dei *targets* molecolari di un fattore trascrizionale, Tbx1, coinvolto nella sindrome di DiGeorge (DGS) utilizzando la tecnologia innovativa del DIGE per uno studio comparativo quantitativo. Tbx1 appartiene alla famiglia dei fattori trascrizionali che possiedono un dominio T-box di legame al DNA ed è una proteina essenziale allo sviluppo del sistema cardiovascolare sia nella classe tassonomica umana che murina. La scoperta di mutazioni puntiformi nel gene TBX1 in pazienti affetti da sindrome di DiGeorge dimostra il coinvolgimento del gene stesso nella patogenesi DGS. Pur essendo chiaro il coinvolgimento di Tbx1 nell'insorgenza della patologia, non sono ancora noti gli effetti delle mutazioni puntiformi studiate sulla macchina cellulare. Obiettivo del progetto di ricerca è stata l'identificazione delle proteine la cui espressione dipende dall'attività del fattore trascrizionale Tbx1 con l'ausilio di metodologie di proteomica di espressione differenziale.
- ✓ Caratterizzazione dell'interattoma della proteina ribosomiale Rps19 mediante approcci sperimentali di proteomica funzionale.  
Analogamente ad altre proteine ribosomiali, Rps19 trasloca dal citoplasma al nucleo, dove partecipa alla biogenesi del ribosoma. E' noto che il 25% di soggetti affetti da DBA presenta mutazioni puntiformi nella sequenza genica che codifica per Rps19. E' stato ipotizzato che le mutazioni trovate incidono sull'espressione proteica e/o determinano una perdita di funzione della proteina stessa, non direttamente connessa alla struttura del ribosoma. Molto resta ancora da chiarire sulla sua funzione non puramente strutturale. La definizione dei ligandi molecolari di nota funzione della proteina può chiarire i processi cellulari in cui la stessa resta coinvolta e far luce sul meccanismo d'azione che la associa alla patogenesi DBA.

## Capitolo 2

### PROCEDURE SPERIMENTALI

#### 2.1 Estrazione di proteine di membrane da RBC

Campioni di sangue umano sono stati prelevati e conservati in una soluzione EDTA 1mg/ml pH 7.4 da soggetti sani (30ml), affetti da HS (10ml) associata a deficit di anchirina/spettina e deficit di band 3. I campioni di sangue di soggetti patologici splenectomizzati e non splenectomizzati sono stati prelevati dalla stessa famiglia.

30 ml di campione di sangue proveniente da soggetti sani (controllo) e 10 ml di campione da soggetti affetti da HS sono stati filtrati [25] e centrifugati a 1000rpm per 10 minuti. Eliminato il sovrantante, gli eritrociti sono stati separati in frazioni di differente densità. Il gradiente di densità è stato realizzato mediante una soluzione Percoll 854g/l, NaKHPO<sub>4</sub> 10mM, NaCl 144mM, EDTA 0.5M, D-glucose 5g/l, PMSF 30µg/ml, osmolalità 320 mOsm/Kg, pH 7.4 [26]. 3ml di eritrociti sono stati risospesi in 35 ml della precedente soluzione, centrifugati a 16000rpm per 35 minuti a 4°C in un tubo di polipropilene, in un rotore ad angolo fisso. Le frazioni di diversa densità sono state raccolte per aspirazione e risospese in un volume, cinque volte superiore di una soluzione NaCl 0.9% p/v, K<sub>2</sub>HPO<sub>4</sub> 5mM pH 8.0 e centrifugati a 1000 rpm per 10 minuti a 4°C. Le cellule sono state lavate centrifugate e quattro volte nella precedente soluzione. Per la preparazione delle membrane eritroidi è stato utilizzato un protocollo convenzionale [26]. Le cellule sono state lisate con K<sub>2</sub>HPO<sub>4</sub> 5mM pH 8.0, contenente PMSF 1mM e EDTA 1mM, e la sospensione è stata centrifugata a 14000rpm per 10 min a 4°C. Scartato il sovrantante le membrane sono state lavate con K<sub>2</sub>HPO<sub>4</sub> 5mM pH 8.0 fino a completa decolorazione. Il campione è stato infine centrifugato a 30000rpm per 1 ora a 4°C in un'ultracentrifuga OPTIMA L-80XD (Beckmann) in un rotore Ti-90.

Il precipitato proteico è stato risospeso in una soluzione urea 7M, tiourea 2M, amidosulfobetaine-14 (ASB-14) 2% p/v, Tris-HCl 40mM pH 8.5.

#### 2.2 Separazione elettroforetica di estratti proteici di membrana, colorazione e analisi d'immagine

Le proteine di membrana sono state ridotte con tributilfosfina (TBP) 3mM per 1ora a 56°C e carboamidometilate con iodoacetamide (IAM) 10mM per 30 min a temperatura ambiente al buio. La soluzione è stata centrifugata a 14000rpm per 30 min; ad essa sono state aggiunte anfoline pH 3-10 1% v/v. Le IPG strips (Biorad), pH 3-10 NL, 17 cm, sono state reidratate attivamente con un voltaggio di 50 V in una soluzione urea 7M, tiourea 2M, amidosulfobetaine-14 (ASB-14) 2% p/v, Tris-HCl 40mM pH 8.5 contenente 60µg o 500µg rispettivamente per i gels analitici e semipreparativi, in una camera Protean IEF Cell (Biorad) per 16 ore. Le strips reidratate sono focalizzate utilizzando le seguenti condizioni: da 0 a 1000V in 5 ore, 1000V costanti per 10 ore, da 1000V a 10000V per un totale di 80KVh. Le strips sono equilibrate per 30 min in un tampone d'equilibratura, contenente urea 6M, SDS 2% p/v, glicerolo 20% v/v, Tris-HCl 0.375M pH 8.5, e caricate su un gel poliaccrilammide/SDS 10%. La corsa elettroforetica è stata condotta a 30mA/gel a 12°C. Le proteine sono state visualizzate utilizzando una doppia procedura di colorazione: AgNO<sub>3</sub>

[27] per i gels analitici e Coomassie colloidale (Pierce) per quelli preparativi. Le immagini relative alle corse eseguite sono state acquisite con il densitometro GS-800 Calibrated Densitometer (Biorad) con una risoluzione 36.3 x 36.3 microns, e analizzate mediante il programma per analisi d'immagine PD-QUEST 7.1 (Biorad). Sono stati studiati tre campioni controllo e sono stati realizzati tre esperimenti 2DE per ogni campione. Gli spots sono stati rivelati e quantizzati automaticamente. Per gli studi comparativi sono stati considerati quegli spots la cui intensità subiva, da un campione all'altro, una variazione pari a 2,5 volte. Gli spots d'interesse sono stati escissi dal gel, sottoposti ad idrolisi *in situ* e successiva analisi per spettrometria di massa MALDI-TOF.

### 2.3 Analisi mediante *Western Blotting* di estratti proteici di membrana

Gli estratti proteici di membrana di cellule RBC sono stati separati mediante elettroforesi monodimensionale su un gel di poliacrilammide/SDS 10% e trasferite su di un filtro di nitrocellulosa (Amersham). I filtri sono stati preincubati in una soluzione di latte 5% (Bio-Rad) in PBS per 2 ore a temperatura ambiente, affinché le proteine del latte potessero saturare tutti i siti d'interazione aspecifici. In seguito i filtri sono stati incubati per 1 ora in una soluzione di latte 1% p/v in PBS contenente Tween-20 0.05% v/v contenente rispettivamente l'anticorpo monoclonale che riconosce in modo specifico l'annexina VII (Santa Cruz), e l'anticorpo che riconosce in modo specifico la  $\beta$ -actina (Sigma). Le proteine sono rivelate con un sistema chemiluminescente che sfrutta l'azione dell'enzima per ossidasi, legata ad un anticorpo secondario anti-mouse (Amersham), che in presenza di un reattivo chemiluminescente catalizza la reazione di ossidazione con sviluppo di luce. Quest'ultima imprime un segnale su una lastra autoradiografica.

### 2.4 Estrazione di proteine da cellule P19

Cloni cellulari P19 stabilmente esprimenti Tbx1 (P19-TBX1-TEV-PA) ed un plasmide di controllo (P19-TEV-PA) sono stati forniti dalla sezione colture cellulari del CEINGE Biotecnologie Avanzate s.c.a.r.l. (Napoli). Sei replicati per le due linee cellulari P19, P19-TBX1-TEV-PA e P19-TEV-PA, sono stati cresciuti a 37°C in un'atmosfera di CO<sub>2</sub> al 5% in un terreno Dulbecco-Modified Minimal Essential Medium (DMEM, Gibco), supplementato con siero fetale bovino (FBS, Gibco) 10% v/v e una miscela 1% v/v di penicillina 50 unità/ml e di streptomicina 50 µg/ml (Hiclone). Per la preparazione degli estratti proteici, le cellule cresciute in monostrato, sono state lavate in PBS freddo centrifugate a 1000 rpm per 5 min e risospese in un 500 µl di tampone di lisi Tris-HCl 30 mM pH 8.5, CHAPS 4% p/v, urea 7M, thiourea 2M, inibitori di proteasi 1X (Amersham). Dopo incubazione a 4°C, per alcuni min, sotto agitazione l'estratto è stato sonicato con un sonicatore ad intermittenza (30 impulsi di lieve intensità), per eliminare contaminanti quali DNA genomico e aggregati lipidici. La sospensione è stata chiarificata mediante centrifugazione a 14000 rpm a 4°C per 30 min in una centrifuga Eppendorf. Il sovrantante è stato precipitato mediante 2D clean up (Amersham) e risospeso nel minimo volume (100µl) di tampone di caricamento CHAPS 4% p/v, urea 7M, tiourea 2M, inibitori di proteasi 1X (Amersham). E' stato necessario rendere basico il pH della miscela proteica in esame (pH 8.0 - 9.0) con una soluzione NaOH 50mM, per agevolare la successiva reazione tra le molecole proteiche e i coloranti fluorescenti. Condizione necessaria alla formazione del legame covalente è, infatti, la presenza del gruppo  $\epsilon$ -amminico delle lisine in forma

deprotonata.

Il contenuto proteico è stato, infine, quantizzato secondo la procedura del 2D Quant Kit (Amersham). La lettura dell'assorbanza dei campioni è stata eseguita a 480nm. Gli estratti proteici ottenuti rispettivamente da cellule P19-TBX1-TEV-PA e P19-TEV-PA sono stati denominati Tbx1 I, Tbx1 II, Tbx1 III, Tbx1 IV, Tbx1 V, Tbx1 VI e controllo I, controllo II, controllo III, controllo IV, controllo V, controllo VI.

## 2.5 Controllo dell'espressione di Tbx1 in cellule P19-TBX1-TEV-PA

Gli estratti proteici di cellule P19-TBX1-TEV-PA e P19-TEV-PA sono stati separati mediante elettroforesi monodimensionale su un gel di poliacrilammide/SDS 10% e trasferite su di un filtro di nitrocellulosa (Amersham). I filtri sono stati preincubati in una soluzione di latte 5% (Bio-rad) in PBS per 2 ore a temperatura ambiente. In seguito i filtri sono stati incubati per 1 ora in una soluzione di latte 1% p/v in PBS contenente Tween-20 0.05% v/v e l'anticorpo secondario anti-IgG (Upstate), nella diluizione 1:10000.

L'eccesso di anticorpo è stato allontanato mediante 3 lavaggi successivi in Tween-20 0.05% in PBS. Le proteine sono state poi rivelate con il sistema della chemiluminescenza ECL (Amersham) utilizzando delle lastre autoradiografiche (Amersham).

## 2.6 Saggi di controllo: *labelling efficiency e same same same*

Per valutare l'efficienza di marcatura del campione con i fluorofori sono state allestite per 30 min a 4°C, al buio, le seguenti reazioni di marcatura:

- ✓ 50µg di estratto proteico totale di cellule P19 sono stati incubati con 400pmol di Cydye Cy5
- ✓ 50µg di estratto proteico di E.Coli, usato come controllo, sono stati incubati con 400pmol di Cydye Cy5

Le reazioni di marcatura sono state bloccate con L-lisina 10mM per 10 min a 4°C al buio. Le miscele proteiche sono state frazionate su un gel poliacrilammide/SDS 10% in quantità decrescenti: 25µg, 12.5µg, 6.25µg. I gels sono stati acquisiti mediante il programma DyCyder e processati con il programma Image Quant (Amersham).

Per determinare l'efficienza di marcatura il volume di fluorescenza relativo al canale del campione in esame è rapportato al volume del canale del campione di controllo, caricato nella medesima quantità. Per escludere la marcatura preferenziale dell'estratto proteico in esame con uno dei tre fluorofori utilizzati, è stato eseguito il seguente esperimento (*Same Same Same*): 50µg del campione in esame sono stati marcati con 400pmol di ogni fluoroforo. La miscela è stata frazionata con una separazione elettroforetica bidimensionale ed il gel è stato acquisito e processato mediante il programma DyCyder.

E' stata calcolata la variazione tra l'abbondanza relativa nelle tre immagini. Poiché l'esperimento è stato allestito con il medesimo campione, i rapporti di abbondanza relativa ( $R = V_1/V_2$ ) dovrebbero essere pari ad 1. Il valore trovato, 1.15, è utilizzato come limite inferiore nell'esperimento comparativo, al di sotto del quale la variazione non è considerata significativa.

## 2.7 Marcatura covalente di estratti proteici da cellule P19 con i fluorofori CyDye DIGE

10 nmol dei fluorofori CyDye Cy2, CyDye Cy3, CyDye Cy5, sono ricostituiti in un volume di dimetilformammide (DMF) tale da ottenere una concentrazione finale pari a 1 nmol/ $\mu$ l. La soluzione ottenuta è stata diluita a 400 pmol/ $\mu$ l in DMF e usata secondo il seguente rapporto: 400 pmol per 50  $\mu$ g di proteina. Sono state allestite per 30 min a 4°C, al buio, le seguenti reazioni di marcatura:

- ✓ 50  $\mu$ g di ogni campione sono stati incubati con 400pmol di CyDye Cy3
- ✓ 50  $\mu$ g di ogni campione sono stati incubati con 400pmol di CyDye Cy5
- ✓ una miscela proteica ottenuta da tutti i campioni in esame (6 estratti da cellule P19-TBX1-TEV-PA e 6 da cellule P19-TEV-PA, rispettivamente campione e controllo) presenti nella stessa quantità (50  $\mu$ g) è stata incubata con 4,8nmol di CyDye Cy2. Quest'ultima rappresenta lo standard interno.

Le reazioni sono state bloccate in presenza di una soluzione di L-lisina 10mM per 10 minuti a 4°C al buio. Le miscele proteiche derivate da cellule esperimenti Tbx1, dai controlli e lo standard interno sono riuniti secondo lo schema riportato in tabella 1. Sono stati allestiti 6 esperimenti di elettroforesi bidimensionale. Con una procedura analoga sono stati realizzati esperimenti con un intervallo di pH più stretto. Lo schema delle miscele di reazione è riportato in tabella 2.

**Tabella 1. Disegno sperimentale per l'analisi differenziale nell'intervallo di pH 3-11**

Gel	Cy3 (400pmol)	Cy5 (400pmol)	Cy2 (400pmol)
1	Tbx1 I (50ug)	Controllo I (50ug)	Pool standard (50ug)
2	Tbx1 II (50ug)	Controllo II (50ug)	Pool standard (50ug)
3	Tbx1 III (50ug)	Controllo III (50ug)	Pool standard (50ug)
4	Controllo IV (50ug)	Tbx1 IV (50ug)	Pool standard (50ug)
5	Controllo V (50ug)	Tbx1 V (50ug)	Pool standard (50ug)
6	Controllo VI (50ug)	Tbx1 VI (50ug)	Pool standard (50ug)

**Tabella 2. Disegno sperimentale per l'analisi differenziale nell'intervallo di pH 4-7**

Gel	Cy3 (400pmol)	Cy5 (400pmol)	Cy2 (400pmol)
1	Tbx1 I (50ug)	Controllo I (50ug)	Pool standard (50ug)
2	Tbx1 II (50ug)	Controllo II (50ug)	Pool standard (50ug)
3	Controllo III (50ug)	Tbx1 III (50ug)	Pool standard (50ug)
5	Controllo IV (50ug)	Tbx1 IV (50ug)	Pool standard (50ug)

## 2.8 Separazione elettroforetica di estratti proteici da cellule P19 e colorazione

I campioni riuniti sono stati diluiti in una soluzione di caricamento urea 7M, tiourea 2M, CHAPS 4% p/v, DTT 130mM contenente una miscela di anfoliti pH 3-10 1% v/v, e caricati rispettivamente su 6 strips d'acrilammide pH 3-11 18cm (Amersham), reidratate con 350 $\mu$ l di una soluzione di reidratazione urea 7M, tiourea 2M, CHAPS 4% p/v, DTT 130mM e anfoliti pH 3-10 0,2% v/v. Il protocollo di isoelettrofocalizzazione (IEF) prevede le seguenti fasi: 300V costanti per 4h, da 300 a 1000V in 6h, da 1000V a 8000V in 3h, 8000v per 1h. Dopo l'IEF le strips sono state equilibrate per 15 minuti in soluzione di equilibratura urea 7M, tiourea 2M, Tris/HCl 100mM, glicerolo 30% v/v, SDS 2% p/v contenente DTT 0,5% p/v e per successivi 15 min in soluzione di equilibratura, contenente IAM 4,5% p/v. La seconda dimensione è stata eseguita su gel poliacrilammide/SDS 11% 18x20cm nel sistema Ettan DALTWELVE (Amersham) applicando 2 W per ogni gel per 16 ore a 15°C. Il medesimo esperimento è stato condotto scegliendo strips pH 4-7 18cm.

Sono stati eseguiti due gels semipreparativi (per estratti da cellule P19-TBX1-TEV-PA e P19-TEV-PA, rispettivamente campione e controllo), per l'esperimento avente pH 3-11 che per quello pH 4-7 con le medesime modalità sia nella prima che nella seconda dimensione. Sono stati utilizzati 500 $\mu$ g di campione non marcato per ognuno dei gel semipreparativi, ottenuti dalla miscela dei corrispondenti replicati biologici. La visualizzazione delle mappe elettroforetiche è stata realizzata incubando i gels prima in 100ml di una soluzione di fissaggio (CH<sub>3</sub>OH 50%, CH<sub>3</sub>COOH 7%) per 3 ore sotto agitazione, quindi in 100ml di una soluzione Sypro Ruby per 16ore sotto agitazione.

## 2.9 Analisi d'immagine mediante il programma DyCyder Differential Analysis

L'acquisizione dei gels analitici ottenuti per l'esperimento con strips di pH 3-11 e di pH 4-7 è stata ottenuta utilizzando uno scanner Typhoon 9400 Variable Mode Imager (Amersham) mediante tre scansioni consecutive per ogni gel, alle seguenti lunghezze d'onda d'eccitazione e di emissione: 532nm/580nm per CyDye Cy3, 633/670nm per CyDye Cy5; 488nm/520nm per CyDye Cy2. I gels preparativi sono stati acquisiti nell'intervallo 457nm/610nm. L'analisi d'immagine è stata eseguita mediante DyCyder DIA e DyCyder BVA impostando i seguenti filtri di esclusione: R>1,30 e R<1,30; P<0.001. Gli spots d'interesse sono stati escissi dai gels preparativi mediante un sistema automatico Ettan Spot Picker (Amersham), sottoposti ad idrolisi *in situ* e analizzati per spettrometria di massa tandem.

## 2.10 Purificazione dei ligandi proteici delle proteine di fusione GST e GST-Rps19

Il cDNA del gene RPS19 è stato amplificato e clonato nel vettore plasmidico pGEX, ed espresso come proteina di fusione GST-Rps19 in *E.Coli*. In modo analogo è stata sintetizzata la proteina GST, usata come controllo. Le proteine GST e GST-Rps19, purificate dai lisati batterici, sono state cimentate con estratti totali di cellule (10<sup>8</sup>) umane eritroleucemiche K562.

La resina glutatione sefarosio è stata equilibrata con 300 µg di proteina GST in PBS/ DTT 1 mM per 1 ora a 4°C sotto agitazione. Le proteine non legate alla resina sono state recuperate mediante centrifugazione a 5.000 rpm per 1 min e incubate con la stessa quantità e lo stesso volume di proteina di fusione GST-Rps19 per 16 ore a 4 °C sotto agitazione. Le proteine non legate sono state allontanate mediante tre lavaggi con 1ml di un tampone Tris HCl 50 mM pH 8, MgCl<sub>2</sub> 12.5 mM, EDTA1 mM, DTT 1 mM, PMSF 1 mM, KCl 0.1M, mentre quelle legate, eluite mediante una soluzione denaturante contenente Tris HCl 50 mM pH 8,12, MgCl<sub>2</sub> 5 mM, EDTA 1 mM, DTT 1 mM, PMSF 1 mM, KCl 0.5M; ed infine precipitate in di F<sub>3</sub>CCOOH 20%.

## 2.11 Analisi per *western blotting* e immunoprecipitazione dei ligandi proteici delle proteine di fusione GST e GST-Rps19

Gli estratti proteici sono stati separati mediante elettroforesi monodimensionale su un gel di poliaccrilammide/SDS 10% e trasferite su di un filtro di nitrocellulosa (Amersham). I filtri sono stati preincubati in una soluzione di latte 5% (Bio-Rad) in PBS per 2 ore a temperatura ambiente. In seguito i filtri sono stati incubati per 1 ora in una soluzione di latte 1% in PBS Tween-20 0.05% con opportune diluizioni degli anticorpi primari

- ✓ anti- DDX5 (Upstate), 1:2000
- ✓ anti-IGF2BP1 (Santa Cruz), 1.1000
- ✓ anti- MCM6 (Santa Cruz), 1.1000
- ✓ anti-STAU1(University of Montreal, Canada), 1.1000
- ✓ anti-PIM1(Santa Cruz), 1.1000

L'eccesso di anticorpo è stato allontanato mediante 3 lavaggi successivi in Tween-20 0.05% in PBS, ed i filtri incubati per 30 minuti con l'opportuno anticorpo secondario coniugato con perossidasi di rafano, diluito nella soluzione di latte all'1%, Tween-20 0.05% in PBS. Le proteine sono state poi rivelate con il sistema della chemiluminescenza ECL (Amersham) utilizzando delle lastre autoradiografiche (Amersham). Le co-immunoprecipitazioni sono state allestite incubando una soluzione Triton X-100 0.5%, contenente l'anticorpo monoclonale anti-Rps19 con 1mg di lisati proteici di cellule K562. I lisati proteici sono stati incubati con una resina di agarosio, derivatizzata con la proteina G (Sigma-Aldrich) per 1 ora a 4°C sotto agitazione. Il sovrantante è stato poi incubato con la resina agarosio/proteinaG/anti-Rps19 per 16 ore a 4°C sotto agitazione. Il medesimo esperimento è stato allestito con una resina agarosio/proteinaG/anti-HA (Santa Cruz Biotechnology) ed usato come controllo negativo. Gli eluati sono stati precipitati e risospesi in un tampone di caricamento, 50 mM Tris HCl pH 6.8, 2% SDS, 10% glicerolo, 100 mM DTT , 0.01% blu di bromofenolo, risolte su gel di poliaccrilammide/SDS 10% e analizzati per western blotting con anticorpi specifici anti

- ✓ DDX5
- ✓ MCM6
- ✓ IMP1
- ✓ PIM e
- ✓ STAU1

## 2.12 Analisi *in silico* degli interattori proteici di Rps19

E' stata realizzata una lista di interattori proteici di Rps19, distinti tra primari se realizzano un'interazione diretta, e secondari un'interazione indiretta, utilizzando la banca dati Human Protein Reference Database ([www.hprd.org](http://www.hprd.org)). Ad aprile 2006 la banca dati conteneva

20,097 specie proteiche, 33,710 interazioni documentate e 171,677 collegamenti alla letteratura (PubMED). E' stata definita una mappa delle interazioni in silico, confrontata con quella realizzata in questo lavoro sperimentale. E' stata, inoltre realizzata una lista di interazioni primarie per ogni proteina identificata.

### **2.13 Separazione elettroforetica dei ligandi delle proteine di fusione GST e GST-Rps19 e colorazione**

I precipitati proteici, ottenuti da esperimenti di cromatografia d'affinità, utilizzando rispettivamente le due esche GST e GST-Rps19 sono stati risospesi in 50µl di tampone di caricamento 50 mM Tris HCl pH 6.8, 2% SDS, 10% glicerolo, 100 mM DTT 0.01% blu di bromofenolo e frazionati mediante elettroforesi monodimensionale su gel poliacrilammide/SDS in gradiente 8-18%. Il pattern elettroforetico delle due miscele proteiche frazionate è stato visualizzato in seguito ad incubazione del gel in 100ml di una soluzione Gel Code Blue Staining Reagent (Pierce) per 16 ore sotto agitazione. Ambedue i canali ottenuti, per GST e GST-Rps19, sono stati sezionati in 65 bande da 2 mm. Ogni banda è stata escissa dal gel e sottoposta ad idrolisi *in situ* e successiva analisi per spettrometria di massa tandem.

### **2.14 Idrolisi *in situ***

Le bande e gli spots d'interesse sono stati escissi dal gel e lavati con due passaggi consecutivi che adoperano rispettivamente acetonitrile (ACN) e  $\text{NH}_4\text{HCO}_3$  0.1M fino a decolorazione. Le proteine presenti nelle bande/spots d'acrilammide sono state ridotte in DTT 10mM,  $\text{NH}_4\text{HCO}_3$  0.1M per 45 min a 56°C, e carboamidometilate in IAM 55mM,  $\text{NH}_4\text{HCO}_3$  0.1M per 30 min a temperatura ambiente al buio. Dopo ulteriori lavaggi in ACN e  $\text{NH}_4\text{HCO}_3$  0.1M, le bande disidratate sono state sottoposte ad idrolisi enzimatica in una soluzione  $\text{NH}_4\text{HCO}_3$  0.05M pH 8.5, contenente tripsina 10ng/µl, per 45 min a 4°C. La soluzione in eccesso è stata rimossa e incubata con una soluzione  $\text{NH}_4\text{HCO}_3$  0.05M pH 8.5 per 16 ore a 37°C. E' stato usato un volume minimo di reazione, sufficiente alla completa idratazione delle bande. Le miscele peptidiche sono state estratte, incubando le bande in acetonitrile per 15 min a temperatura ambiente e liofilizzate.

### **2.15 Spettrometria di massa MALDI TOF e analisi dei dati**

I digeriti triptici sono stati liofilizzati e risospesi in una soluzione  $\text{F}_3\text{CCOOH}$  0.5% v/v. La miscela peptidica è stata desalificata mediante minicromatografia in fase inversa con zip tip  $\text{C}_{18}$  (Millipore). Il campione è stato co-cristallizzato con una matrice organica, acido  $\alpha$ -ciano-4-idrossi-cinnamico per la successiva analisi condotta con uno strumento MALDI-TOF Voyager-DE PRO (Applied Biosystems). Gli spettri massa sono stati acquisiti con i seguenti parametri: modalità positiva, modalità reflectron, potenziale di accelerazione 20000V, voltaggio griglia 70%, tempo di estrazione 100ns, intensità del laser 2000 e accumulo di 1000 acquisizioni per ogni spettro. I peptidi che derivano dall'autodigestione triptica (842.51 2163.06, 2274.46, m/z) sono stati utilizzati come calibranti interni. Per ogni miscela peptidica è stato ottenuto uno spettro di massa e una lista di segnali, corrispondenti ad i valori  $\text{MH}^+$  di ogni peptide.

Le proteine sono state identificate utilizzando il programma *Mascot* ([www.matrixscience.com](http://www.matrixscience.com)) imponendo i seguenti parametri di ricerca:

- ✓ banca dati proteica non ridondante, NCBI nr
- ✓ enzima proteolitico specifico, tripsina
- ✓ classe tassonomica del campione, umana
- ✓ un unico sito d'idrolisi non riconosciuto dall'enzima lungo la sequenza
- ✓ presenza di S- carboamidometilcisteine
- ✓ presenza di estremità N e C terminali non modificate
- ✓ presenza di metionine ossidate e non ossidate
- ✓ possibile formazione del piroglutammico all'estremità N-terminale
- ✓ tolleranza pari ad un valore massimo di 150 ppm

I risultati sono stati valutati in base al numero di peptidi corrispondenti tra lo spettro e la proteina presente in banca dati, il pl, il peso molecolare apparente ottenuto dalla migrazione elettroforetica, la copertura di sequenza. In alcuni casi la miscela peptidica è stata analizzata mediante spettrometria di massa tandem.

## 2.16 Spettrometria di massa tandem e analisi dei dati

Le miscele peptidiche liofilizzate sono state solubilizzate in 20µl di una soluzione F<sub>3</sub>CCOOH 0.5%. L'analisi per spettrometria di massa è stata condotta mediante µLCMS/MS utilizzando due differenti strumenti: uno spettrometro di massa ibrido Q-ToF (Waters) dotato di una sorgente electrospray e accoppiato con un sistema cromatografico *capLC* (Waters), ed un LC/MSD Trap XCT Ultra (Agilent Technologies) dotato di un sistema HPLC 1100 e di un *chip cube* (Agilent Technologies). La miscela peptidica caricata (7µl) è stata concentrata e desalificata (i) ad 1µl/min in una precolonna C<sub>18</sub> a fase inversa (Waters) oppure (ii) a 4µl/min in una colonna di arricchimento a fase inversa (Agilent Technologies) con un eluente HCOOH 0.1% v/v. Il campione è stato frazionato su una colonna cromatografia capillare C<sub>18</sub> a fase inversa (le cui dimensioni nel sistema Waters sono 75µm x 20 cm, nel sistema Agilent 75µm x 43mm) ad una velocità di flusso pari a 200nl/min con un gradiente lineare di eluente B (HCOOH 0.1% v/v in acetonitrile) nell'eluente A (HCOOH 0.1% v/v), dal 5 al 60% in 50 min. Lo strumento esegue l'analisi acquisendo spettri di massa delle frazioni eluite dal sistema cromatografico in un intervallo tra 400 e 2000m/z, secondo la modalità *data-dependent*. Infine acquisisce gli spettri MS/MS dei tre ioni più abbondanti presenti in ogni spettro MS. E' stata utilizzata un'energia di collisione tra 30 e 60 eV in dipendenza dalla massa e dallo stato di carica dello ione precursore da frammentare. Per evitare l'acquisizione di spettri MS/MS di peptidi dovuti a cheratine e/o alla tripsina stessa il metodo d'acquisizione include una lista d'esclusione dei più frequenti segnali contaminanti recanti doppia e tripla carica: 403,20; 517,00; 519,32; 525,00; 532,90; 559,32; 577,30; 587,86; 616,85; 721,75; 745,90; 747,32; 758,43; 854,30; 858,43; 896,30; 1082,06.

I *files* contenenti le liste peptidiche, ottenute dagli spettri di massa, sono stati convertiti in un formato di testo mediante il programma ProteinLynx Global Server al fine di eseguire le identificazioni proteiche con il programma *MASCOT* (Matrix Science; [www.matrixscience.com](http://www.matrixscience.com)). Le proteine sono state identificate imponendo i seguenti parametri di ricerca:

- ✓ banca dati proteica non ridondante, NCBI nr
- ✓ enzima proteolitico specifico, tripsina
- ✓ classe tassonomica del campione umana
- ✓ un unico sito d'idrolisi non riconosciuto dall'enzima lungo la sequenza

- ✓ presenza di S- carboamidometilcisteine
- ✓ presenza di estremità N e C terminali non modificate
- ✓ presenza di metionine ossidate e non ossidate
- ✓ possibile formazione del piroglutammico all'estremità N-terminale
- ✓ tolleranza pari ad un valore massimo di 150 ppm per la massa del peptide precursore
- ✓ tolleranza pari ad un valore massimo di 300 ppm per la massa del peptide frammento.

Tutte le identificazioni realizzate con uno *Mascot Score* superiore a 38 sono state ritenute non ambigue. Il *Mascot Score* è definito come  $-10 \times \log P$  dove P è la probabilità assoluta che l'evento di confronto sia casuale. Minore è il valore di P maggiore è il valore del *Mascot Score*, più significativa è l'identificazione realizzata.

## Capitolo 3

# CARATTERIZZAZIONE DELLE PROTEINE DI MEMBRANA DEL GLOBULO ROSSO, IN FUNZIONE DELLA DENSITÀ CELLULARE

### 3.1 Introduzione

Il principale ruolo dell'eritrocita è il trasporto dell'ossigeno in risposta alle diverse esigenze dei tessuti. L'eritrocita manifesta le sue funzioni grazie all'interazione di tre elementi, quali la membrana eritrocitaria, l'emoglobina e le vie metaboliche che consentono di regolare il trasporto di ossigeno, proteggere l'emoglobina dai danni ossidativi e preservare la stabilità osmotica dell'ambiente cellulare. L'organizzazione strutturale e la composizione biochimica sono determinanti per la funzionalità della membrana eritroide; da esse dipendono, infatti, parametri cellulari, quali la forma, il volume cellulare medio, la resistenza osmotica e la deformabilità della cellula stessa.

La membrana eritroide è costituita da un doppio strato lipidico e da proteine integrali legate ad uno scheletro proteico sottostante [30]. Il doppio strato lipidico costituisce circa il 50% del peso della membrana ed ha una conformazione asimmetrica: i fosfolipidi colinici sono, infatti, localizzati sulla superficie esterna della cellula, mentre gli aminofosfolipidi sono concentrati sulla superficie interna. Tale conformazione asimmetrica è mantenuta da un sistema di trasporto (flippasi) che è ATP-dipendente. Il "core idrofobico" è attraversato da alcune proteine (proteine integrali di membrana) che costituiscono le vie di comunicazione tra l'ambiente idrofilo intracellulare e quello extracellulare. Tali proteine comprendono: la banda 3, i canali per il trasporto del Na, Na-K-ATPasi e le famiglie delle glicoforine (A, B, C, D). Due sono i punti di attacco del doppio strato lipidico con il citoscheletro. Il primo è costituito dall'interazione tra beta spectrina-proteina 4.1-actina e glicoforina C. Il secondo punto di legame è costituito dall'interazione tra beta spectrina-anchirina-banda 3, a cui partecipa la proteina 4.2 [31]. Il citoscheletro riveste un ruolo importante anche nel trasporto degli ioni, mediato da alcuni canali di membrana o da pompe che permettono di mantenere il gradiente osmotico all'interno della cellula. I sistemi di trasporto cationico di membrana possono essere suddivisi in 3 gruppi secondo le loro caratteristiche funzionali: i trasporti energia-dipendenti come la pompa Na-K-ATPasi, i sistemi gradiente-dipendente quali il cotrasporto K-Cl ed infine i canali ionici come il canale Ca-K attivato, noto come canale Gardos [32-33]. Difetti nelle interazioni proteina-proteina lungo la membrana, o tra membrana e citoscheletro, o nella composizione proteica di membrana, determinano anomalie nei rapporti superficie-volume e nelle concentrazioni dei cationi monovalenti, destabilizzando il doppio strato lipidico con conseguente perdita del materiale proteico. La cellula, pertanto, assume una forma sferica che modifica le sue peculiari proprietà di plasticità e deformabilità. Le anomalie elencate costituiscono le basi molecolari di un'ampia classe di patologie, note come sferocitosi [34-35]. Fenotipi differenti caratterizzano le sferocitosi ereditarie (HS). Possono essere causate da mutazioni genetiche familiari dominanti e recessive. Queste mutazioni interessano la sequenza di cinque proteine che sono coinvolte nelle "interazioni verticali" tra citoscheletro e doppio strato lipidico [36]:  $\alpha$  e  $\beta$  spectrina [32], anchirina [33], banda 3 [37-38], e proteina 4.2 [39]. Le più comuni forme della patologia presentano basse concentrazioni rispettivamente di

una coppia di proteine del citoscheletro, spettrina e anchirina, e di una proteina di membrana, banda 3 [40].

Una riduzione dei sintomi patologici delle sferocitosi per deficit di anchirina e spettrina o per deficit di banda 3 è stata osservata in seguito alla rimozione della milza, un organo preposto alla maturazione di cellule eritroidi immature (neociti) a cellule mature (gerociti).

La complessa fisiopatologia delle patologie che derivano da anomalie riscontrate negli eritrociti può essere chiarita solo attraverso lo studio approfondito della struttura e delle funzioni del globulo rosso normale, e della regolazione dell'eritropoiesi. I globuli rossi umani rappresentano l'ultimo stadio della maturazione eritroide. Durante la loro vita media di 120 giorni mostrano una variazione nell'organizzazione strutturale.

Negli ultimi trent'anni gli studi genetici hanno prodotto numerosi risultati che hanno permesso di descrivere purtroppo in maniera non esauriente le funzioni e la struttura del globulo rosso. Tuttavia, la notevole diversità fenotipica che caratterizza le più comuni patologie del globulo rosso, evidenzia che molti aspetti devono ancora essere chiariti soprattutto per quello che riguarda i risultati ottenuti a livello molecolare, in relazione ai fenotipi clinici [41]. Il problema relativo a qualunque tipo di studio genetico è l'enorme complessità del sistema in esame. Il fenotipo è, infatti, influenzato da numerosi fattori genetici, inclusi i geni che codificano per le proteine coinvolte nel metabolismo dei globuli rossi e nella struttura della membrana eritroide. Per poter studiare l'organizzazione strutturale, la composizione e l'evoluzione temporale della membrana eritroide è stato utilizzato un approccio di proteomica comparativa. Sono stati confrontati i proteomi di due popolazioni cellulari: cellule eritroidi immature (neociti) ed adulte (gerociti). Queste ultime, infatti, pur derivando dal medesimo campione, differiscono per la diversa densità, deformabilità e comportamento in processi cellulari, quali ad esempio la vescicolazione.

Le due frazioni cellulari sono state separate con un gradiente di densità [42] e sono stati analizzati i profili d'espressione delle proteine di membrana estratte dalle frazioni cellulari di differente stratificazione. Lo studio è stato condotto su cellule di soggetti sani e affetti da HS, splenectomizzati e non splenectomizzati. L'approccio proteomico scelto prevede la separazione delle proteine di membrana lungo una mappa elettroforetica bidimensionale, la successiva analisi d'immagine e l'identificazione per spettrometria di massa degli spots d'interesse. Le maggiori difficoltà per l'analisi bidimensionale delle proteine di membrana derivano dalla scarsa efficienza della maggior parte dei comuni detergenti utilizzati al fine di solubilizzare proteine di membrana denaturate.

### 3.1.2 L'approccio proteomico allo studio delle proteine di membrana

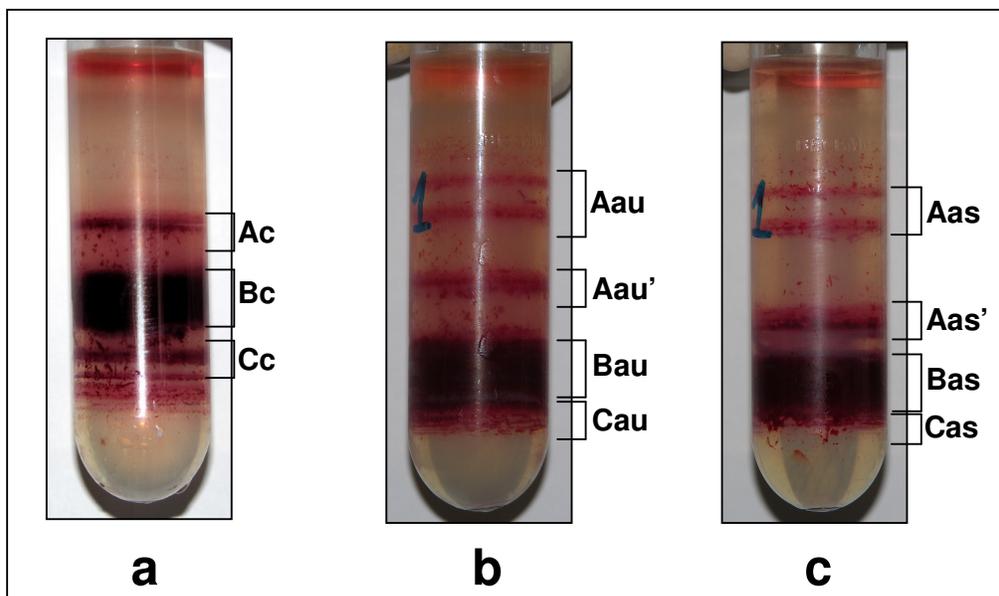
L'analisi delle proteine di membrana attraverso approcci di tipo proteomico è stata, negli anni, oggetto di approfonditi studi. Data la loro locazione sulla superficie cellulare, le proteine di membrana svolgono importanti ruoli in vari processi cellulari, quali la trasduzione del segnale, l'adesione cellulare, il trasporto di metaboliti e ioni, l'endocitosi, etc. Il recente completamento del progetto di sequenziamento del genoma umano mostra che le proteine transmembrana rappresentano il 30% delle proteine totali [43]. Essendo cruciali alla vita cellulare, non è sorprendente l'attenzione scientifica ad esse rivolta. Un'indagine di tipo proteomico-differenziale fonda su tecniche elettroforetiche bidimensionali, che sono di difficile impiego a causa della scarsa solubilità delle proteine di membrana nei convenzionali tamponi acquosi [44]. E' necessario l'uso di detergenti che mimino l'ambiente lipidico interno delle micelle. Per ottenere una mappa bidimensionale risolta è necessario, infatti: i) eliminare la matrice lipidica dal campione in esame per minimizzare le possibili interferenze con la IEF; ii) estrarre le proteine dalla matrice lipidica

convertendole in una forma solubile (complesso proteina-detergente) iii) conservare suddetta solubilità durante la IEF, soprattutto in prossimità dei pI, punti di minima solubilità proteica. Tamponi contenenti convenzionali agenti caotropici, quali urea e tiourea non sono sufficienti a soddisfare le precedenti condizioni. E' stato ipotizzato che la denaturazione proteica dovuta ad agenti caotropici determina la formazione di forti legami nel cuore idrofobico della proteina di membrana aumentando i siti di aggregazione interna, inficiandone la solubilizzazione. Una soluzione al problema risiede nell'introduzione di nuovi detergenti zwitterionici, le ammidosulfobetaine. Queste ultime sono caratterizzate da una coda idrofobica, una testa idrofilica ed una regione di legame tra le precedenti. Il surfattante d'elezione per la solubilizzazione di proteine di membrana è l'ammidosulfobetaina ASB-14, caratterizzato da un gruppo ammidico, un azoto quaternario ed un gruppo solforico, ed una catena di gruppi metilenici [45]. La sua efficienza è stata valutata con la solubilizzazione della banda 3, una proteina che consta di 13 domini transmembrana, non visualizzata nelle convenzionali mappe elettroforetiche bidimensionali.

### 3.2 Risultati

L'organizzazione strutturale, la composizione e l'evoluzione temporale della membrana eritroide sono stati studiati con un approccio di proteomica comparativa. Sono stati confrontati i proteomi di due popolazioni cellulari immature (neociti) ed adulte (gerociti). Queste ultime, infatti, pur derivando dal medesimo campione, differiscono per la diversa densità, deformabilità e comportamento in processi cellulari, quali ad esempio la vescicolazione.

Le due frazioni cellulari provenienti da un soggetto sano sono state separate con un gradiente di densità come mostrato in figura 7. Le frazioni caratterizzate da differente densità, indicate come Ac, Bc e Cc sono state raccolte e processate. Le proteine di membrana relative alle frazioni Ac e Bc sono state analizzate mediante elettroforesi bidimensionale.



**Figura 7.** Distribuzione di cellule di globulo rosso su un gradiente di densità Percoll di un campione sano (a), patologici per HS con deficit di anchirina/spettrina non splenectomizzato (b) e splenectomizzato (c). A, B e C indicano le frazioni raccolte; i pedici 'c', 'as' e 'au' indicano rispettivamente controllo, HS per deficit di anchirina/spettrina splenectomizzati e HS per deficit di anchirina/spettrina non splenectomizzati.

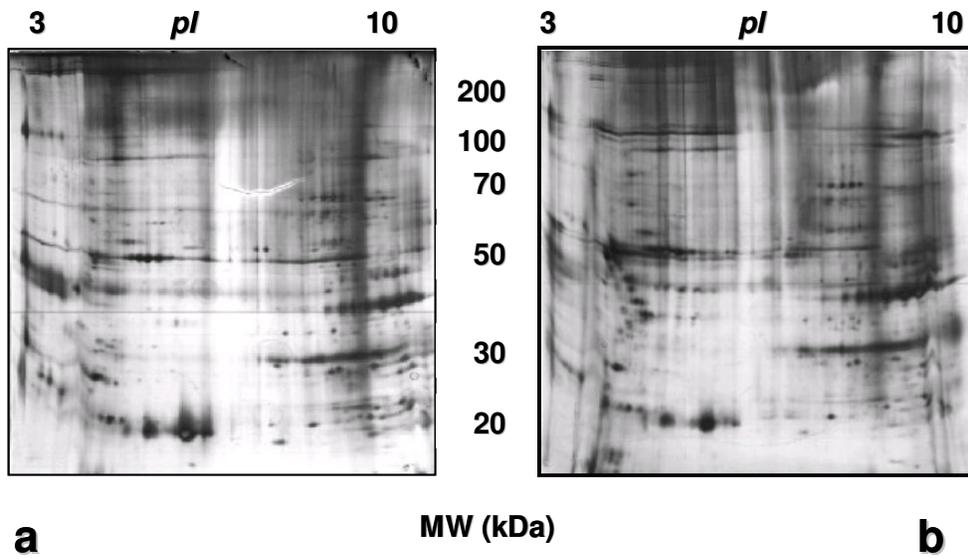
Le maggiori difficoltà per l'analisi bidimensionale delle proteine di membrana derivano dalla scarsa efficienza della maggior parte dei comuni detergenti utilizzati al fine di solubilizzare proteine di membrana denaturate. Sono stati testati diversi agenti surfattanti, in particolare un detergente zwitterionico, caratterizzato da una coda idrofobica, una testa idrofila ed una regione di legame tra le precedenti, ASB-14 (Sigma). L'efficienza, dei detergenti testati, è stata valutata dalla solubilizzazione della "banda 3", una proteina costituita da 13 domini trans-membrana. La figura 8 mostra due delle mappe bidimensionali analitiche realizzate dalle due frazioni, dopo colorazione all'argento. I parametri per l'acquisizione dei gels, realizzata mediante PD QUEST 7.1 sono stati

ottimizzati su una delle immagini per minimizzare la rivelazione di artefatti a massimizzare la rivelazione degli spots. I parametri ottenuti sono stati poi utilizzati per tutte le immagini da analizzare. Gli spots presenti nei tre replicati biologici di ogni campione costituiscono un'immagine composita. Questa immagine composita, ottenuta per i due campioni, è stata usata per il confronto. Per valutare la variazione tra gels abbiamo stimato l'intensità totale (Spot Quantity) di un definito spot (relativo all'actina) in tutti i gruppi di replicati, analitici e biologici (tabella 3). Il coefficiente di variazione tra gels è inferiore al 5% per tutti gli spots analizzati. Questo dato indica una buona riproducibilità del metodo.

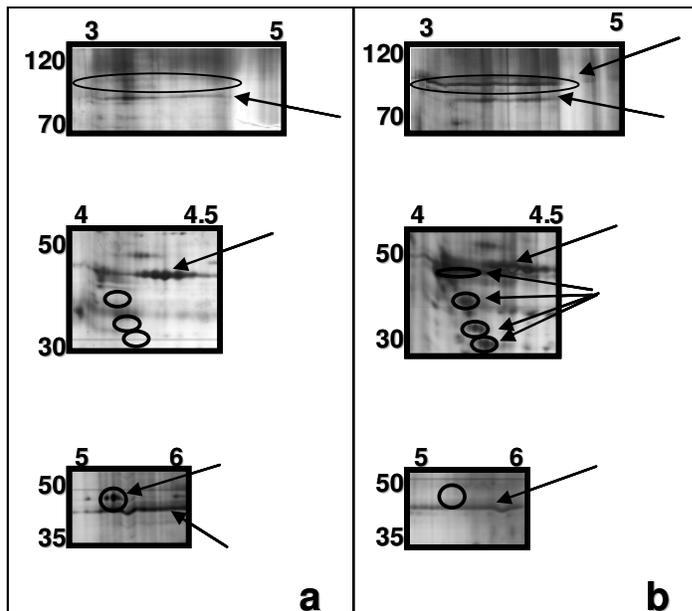
Sono stati visualizzati circa 500 spots proteici. Quelli d'interesse sono stati processati ed identificati mediante la strategia del *peptide mass fingerprinting* con uno spettrometro MALDI-TOF. Come mostrato dalla figura 9, sono state individuate sulle mappe elettroforetiche delle due popolazioni alcune differenze. Le proteine relative agli spots differenza tra i due campioni sono elencate nella tabella 4. L'isoforma dell'anchirina R con peso molecolare compreso tra 80 e 100KDa è più espressa nella frazione meno densa. Nella frazione più densa sono invece prodotti della degradazione proteolitica del frammento citosolico della banda 3. Non avendo osservato i medesimi frammenti anche nella frazione meno densa possiamo confermare che la degradazione è dovuta all'invecchiamento cellulare e non alla procedura sperimentale. Infine uno spot proteico di peso molecolare compreso tra 50 e 45 KDa e di punto isoelettrico tra 5 e 6 è presente nella frazione Ac, ma non nella Cc. La proteina è stata identificata come annexina VII mediante il programma Mascot con uno score pari a 223. La miscela peptidica è stata anche processata per LCMS/MS come descritto nella procedura sperimentale. I peptidi identificati derivano da regioni presenti sull'intera sequenza della proteina restituendo uno score pari a 531.

Solo recentemente è stato riportato che l'annexina VII fa parte dei globuli rossi [46]. La figura 10 mostra che l'annexina è presente solo nella frazione Ac, mentre il segnale è assente nelle altre frazioni analizzate: Bc e Cc. Inoltre sono state utilizzate diverse quantità di proteine di membrana per definire il limite inferiore di rilevabilità. La minima quantità proteica necessaria alla visualizzazione del segnale relativo alla presenza di annexina VII è pari a 100 µg della frazione Ac. Le quantità proteiche utilizzate sono state normalizzate eseguendo un *western blotting* di tutte le frazioni (10µg) analizzate contro l'actina: è possibile visualizzare il segnale relativo alla presenza di actina in tutte le frazioni. Il risultato conferma la perdita di annexina VII durante la maturazione dei reticolociti.

L'espressione di annexina VII è stata studiata in campioni patologici per HS splenectomizzati e non splenectomizzati, per valutare la responsabilità della milza nella patologia [42]. Per l'esperimento descritto sono stati utilizzati campioni di sangue provenienti da pazienti affetti da HS per deficit di anchirina e spettrina o per deficit di banda 3. Le figure 7b e 7c mostrano i gradienti di densità dei campioni patologici per deficit di anchirina rispettivamente splenectomizzati e non splenectomizzati. Il profilo del gradiente si mostra profondamente diverso dal controllo: sono state studiate quattro differenti frazioni. Suddette frazioni sono state processate in modo analogo al campione sano, ma la quantità proteica estratta dalle membrane si è dimostrata insufficiente per un'indagine 2DE. Pertanto l'espressione dell'annexina VII è stata seguita mediante esperimenti di *western blotting*. La figura 11a mostra che l'annexina VII è presente solo nelle frazioni più leggere di ambedue i campioni di HS per deficit di anchirina e spettrina studiati.



**Figura 8.** Estratti proteici di membrana delle frazioni cellulari di differente densità Ac (a) e Cc (b) sono stati frazionati su 2DE e colorati con  $\text{AgNO}_3$ .



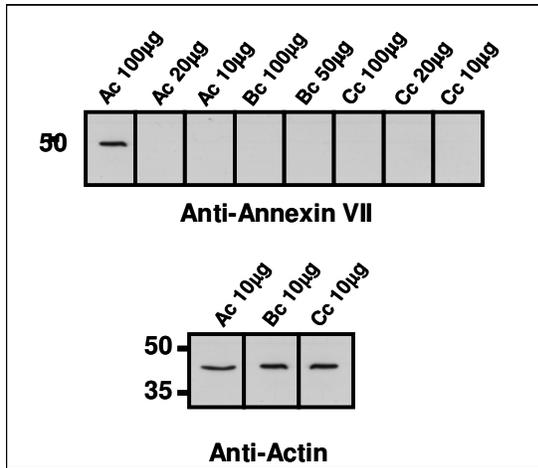
**Figura 9.** In figura sono mostrate le differenze trovate nelle mappe elettroforetiche delle frazioni cellulari di differente densità Ac (a) e Cc (b).

**Tabella 3. Determinazione dello "Spot quantity" dell'actina e misura del CV dei replicati**

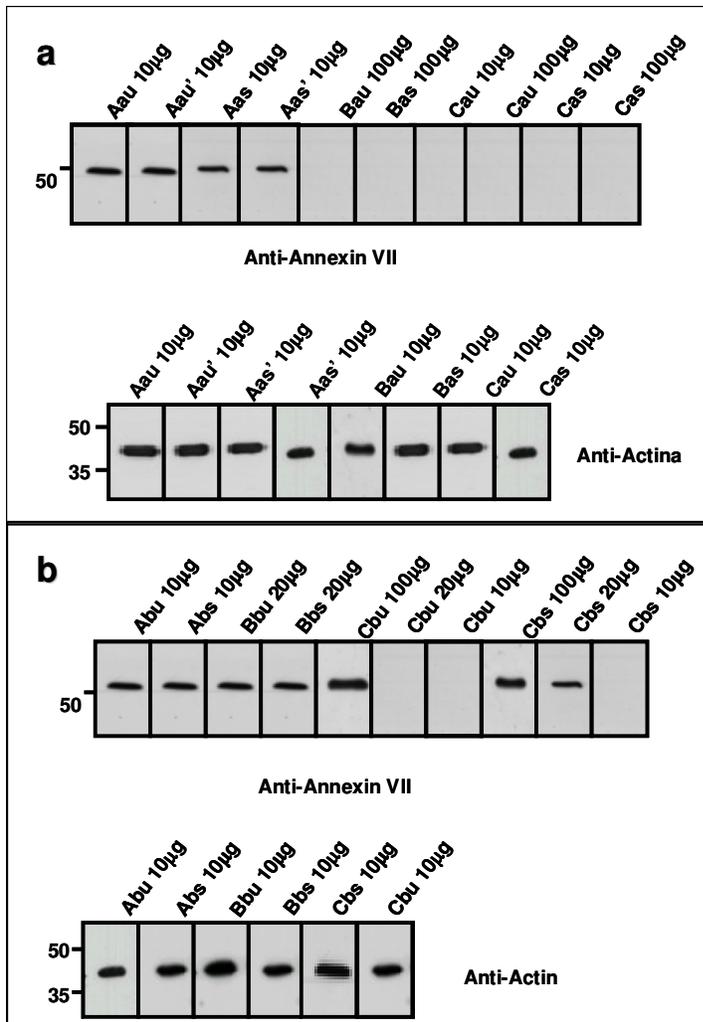
Campione	Frazione cellulare	Spot quantity	Frazione cellulare	Spot quantity
controllo 1	Fraction C (gel 1)	1009	Fraction A (gel 1)	980
	Fraction C (gel 2)	915	Fraction A (gel 2)	1062
	Fraction C (gel 3)	995	Fraction A (gel 3)	1045
controllo 2	Fraction C (gel 1)	1000	Fraction A (gel 1)	987
	Fraction C (gel 2)	919	Fraction A (gel 2)	1050
	Fraction C (gel 3)	999	Fraction A (gel 3)	1035
controllo 3	Fraction C (gel 1)	1020	Fraction A (gel 1)	987
	Fraction C (gel 2)	939	Fraction A (gel 2)	1020
	Fraction C (gel 3)	989	Fraction A (gel 3)	1025
controllo 4	Fraction C (gel 1)	989	Fraction A (gel 1)	1008
	Fraction C (gel 2)	965	Fraction A (gel 2)	1042
	Fraction C (gel 3)	945	Fraction A (gel 3)	1025
controllo 5	Fraction C (gel 1)	1021	Fraction A (gel 1)	1007
	Fraction C (gel 2)	974	Fraction A (gel 2)	984
	Fraction C (gel 3)	992	Fraction A (gel 3)	1006
controllo 6	Fraction C (gel 1)	956	Fraction A (gel 1)	991
	Fraction C (gel 2)	987	Fraction A (gel 2)	1062
	Fraction C (gel 3)	1008	Fraction A (gel 3)	1054
controllo 7	Fraction C (gel 1)	994	Fraction A (gel 1)	1003
	Fraction C (gel 2)	1015	Fraction A (gel 2)	997
	Fraction C (gel 3)	984	Fraction A (gel 3)	995
controllo 8	Fraction C (gel 1)	996	Fraction A (gel 1)	1005
	Fraction C (gel 2)	1011	Fraction A (gel 2)	1087
	Fraction C (gel 3)	1006	Fraction A (gel 3)	1013
controllo 9	Fraction C (gel 1)	1000	Fraction A (gel 1)	993
	Fraction C (gel 2)	995	Fraction A (gel 2)	998
	Fraction C (gel 3)	993	Fraction A (gel 3)	1008
controllo 10	Fraction C (gel 1)	998	Fraction A (gel 1)	996
	Fraction C (gel 2)	989	Fraction A (gel 2)	1015
	Fraction C (gel 3)	991	Fraction A (gel 3)	987
average		986,5		1015,6
SD		27,1		27,6
CV		0,027		0,027

**Tabella 4. Proteine differenzialmente espresse in popolazioni eritroidi di differente età cellulare**

Gene	Proteina	MW (Da)	pI	Copertura di sequenza	MASCOT Score
ANXA7	Annexin VII	47100	5,52	34%	223
BND3	Band 3 anion transport protein (Chain P)	42623	4,72	32%	90
ANK1	Ankyrin	189000	5,72	15%	107



**Figura 10.** L'annexina VII è espressa in modo differente nelle frazioni cellulari di differente densità. Sono necessari 100 µg di estratto proteico totale per la rivelazione dell'annexina VII. 10 µg di estratto proteico sono sufficienti per la rivelazione della β-actina nell'esperimento di controllo.



**Figura 11.** L'annexina VII è espressa in modo differente nelle frazioni cellulari di differente densità di campioni HS per deficit di anchirina/spettrina (a) e banda 3 (b). Sono sufficienti 10µg di estratto proteico totale per la rivelazione dell'annexina VII nelle frazioni leggere di HS per deficit di anchirina 7(a). Nei campioni di HS per deficit di banda 3 il segnale relativo all'annexina VII è presente in tutte le frazioni (b). 10 µg di estratto proteico sono utilizzati per la rivelazione della β-actina negli esperimenti di controllo.

Come per il campione controllo, è stato eseguito un esperimento di western blotting parallelo anti-actina per normalizzare le quantità proteiche delle differenti frazioni analizzate. Il risultato ha evidenziato per il campione patologico studiato un comportamento analogo al controllo. E' stato allestito un secondo esperimento per l'analisi di campioni patologici HS per deficit di banda 3. Anche in questo caso la distribuzione delle cellule di globulo rosso lungo un gradiente di densità si mostra molto diversa dal controllo. Essendo la quantità proteica recuperata insufficiente per uno studio 2DE, l'espressione dell'annexina VII è stata seguita mediante western blotting (figura 11b). E' stato trovato un comportamento profondamente diverso dai precedenti campioni analizzati. Al segnale che rivela la presenza di annexina VII è più intenso nella frazione pesante dei pazienti splenectomizzati. Il limite inferiore di quantità proteica da utilizzare per la visualizzazione del segnale ammonta a 20µg per i pazienti splenectomizzati e a 100µg per quelli non splenectomizzati.

### 3.3 Discussione

Lo studio condotto mostra come un approccio di tipo proteomico può essere usato per rivelare relazioni struttura funzione anomale in difetti di membrana di cellule di globulo rosso.

E' stato confrontato il profilo di espressione delle proteine di membrana dei globuli rossi in campioni differenti provenienti da cellule di età differente in soggetti sani e patologici per HS, mediante un approccio proteomico. L'analisi delle proteine di membrana rappresenta un obiettivo difficile da perseguire per le indagini di tipo proteomico. In particolare studi proteomici, che fondano sull'utilizzo della tecnica 2DE, possono omettere completamente le proteine presenti in basse quantità [47-49]. Tenendo conto di queste limitazioni è stato messo a punto un protocollo sperimentale per lo studio delle proteine di membrana di globuli rossi seguito da uno studio, focalizzato sulle differenze d'espressione proteica di frazioni cellulari di differente densità. I risultati ottenuti dimostrano che l'invecchiamento delle cellule di globulo rosso determina l'accumulo di prodotti dovuti alla proteolisi di proteine del citoscheletro quali anchirina e banda 3. L'anchirina è espressa in varie isoforme tessuto-specifiche. Negli eritrociti umani è espressa l'anchirina R, che presenta un peso molecolare pari a 206 KDa. Le isoforme di peso molecolare 186, 170, e 145 KDa sono presenti in minore quantità e derivano da splicing alternativo o da frammenti proteolitici (170 KDa) di anchirine di peso molecolare più alto [50]. I risultati ottenuti mostrano che l'anchirina è oggetto di modificazioni proteolitiche durante la maturazione. In cellule mature è noto che la banda 3 è oggetto dall'azione proteolitica della caspasi 3 [51], attività silente nelle cellule più giovani.

Il confronto tra le mappe bidimensionali delle proteine di membrana di cellule di globulo rosso con differente stratificazione su un gradiente di densità ha mostrato che la proteina annexina VII è presente nei reticolociti ed è persa durante la maturazione cellulare. I reticolociti hanno una maggiore deformabilità e capacità di vescicolazione mediante la formazione di micro e nano-vescicole rispetto a cellule adulte. Queste proprietà sono legate alla presenza di annexina VII [52]. L'annexina VII è una proteina legata ai fosfolipidi e agli ioni  $Ca^{2+}$ . Appartiene alla famiglia di proteine caratterizzate da una struttura bipolare con un dominio N-terminale variabile ed un dominio C-terminale conservato, responsabile del legame alla membrana lipidica attraverso ioni  $Ca^{2+}$  [49]. Questa proteina è presente nel citosol, sulla membrana plasmatica, intorno ai nuclei e in strutture vescicolari. Trasloca sulla membrana in dipendenza dalla concentrazione intracellulare di  $Ca^{2+}$ . In particolare quando le concentrazioni di  $Ca^{2+}$  aumentano nella cellula la proteina è legata alla membrana e alle vescicole intracellulari. Tuttavia la sua funzione non è ancora nota. Poiché la traslocazione dell'annexina VII è legata all'aumento della concentrazione degli ioni  $Ca^{2+}$  è sorprendente la sua assenza nelle frazioni cellulari pesanti, dove i livelli di  $Ca^{2+}$  sono maggiori. E' possibile che la perdita di annexina VII sia dovuta alla sua degradazione durante l'invecchiamento cellulare o al rimodellamento cellulare per esocitosi. Non sono stati osservati prodotti di degradazione proteolitica dell'annexina VII nelle mappe bidimensionali analizzate. Per valutare il ruolo dell'annexina VII nel rimodellamento cellulare sono stati studiati campioni provenienti da soggetti affetti da HS per deficit di anchirina e spettina o per deficit di banda 3.

La patologia per deficit di anchirina e spettina determina un rimodellamento del citoscheletro [40]. Detto rimodellamento non si osserva nella patologia per deficit di banda. Infatti nel deficit di banda 3 l'annexina VII è presente durante l'intera vita cellulare mentre nel deficit per anchirina e per spettina l'annexina è persa durante l'invecchiamento, come nel controllo. Nel deficit di banda 3, la vescicolazione coinvolge le proteine integrali di membrane e quelle del citoscheletro. Questo potrebbe spiegare la presenza dell'annexina

VII nelle cellule più adulte. Infine la milza non sembra esercitare alcun effetto sulla rimozione di annexina VII, in quanto la presenza di questa proteina in cellule mature mostra che campioni splenectomizzati e non splenectomizzati hanno un comportamento analogo.

## Capitolo 4

# CARATTERIZZAZIONE DEI TARGET MOLECOLARI DEL FATTORE TRASCRIZIONALE *Tbx1*. SINDROME DI DIGEORGE

### 4.1 Introduzione

La sindrome di DiGeorge (DGS) è un esempio di disordine nello sviluppo dell'apparato faringeo. Molti casi di DGS sono associati a delezioni cromosomiche eterozigotiche, che interessano il cromosoma 22q11.2. Sono tra le più frequenti delezioni cromosomiche associate alla nascita (1 ogni 4000 nati vivi), e si manifestano con un'ampia variabilità fenotipica. Esperimenti condotti su modelli murini mostrano che delezioni eterozigotiche lungo una regione cromosomica di 1 Mb, omologa a quella umana, determinano, nell'animale, anomalie cardiovascolari caratteristiche della patologia umana. L'unico gene aploinsufficiente nella regione cromosomica studiata è *TBX1* [53]. L'omologo umano è incluso nel cromosoma 22q11.2. La presenza di mutazioni puntiformi in molti pazienti con fenotipi caratteristici della DGS dimostra in modo conclusivo il ruolo del gene nella patogenesi. I geni *T-BOX* codificano per una famiglia di proteine DNA-binding filogeneticamente conservate che regolano l'espressione genica durante l'embriogenesi [54-56]. Le proteine che recano un dominio T-box legano sequenze nucleotidiche palindromiche, *T-box binding element* (TBE) come dimeri o come monomeri [57-58]. Inoltre hanno la capacità di legare altri fattori trascrizionali, attivando in modo cooperativo la trascrizione dei geni bersaglio [63]. *TBX1* è l'unico gene richiesto per lo sviluppo di tutte le strutture e gli organi che derivano dall'apparato faringeo [60-61]. *TBX1* è espresso nell'endoderma faringeo, nel mesoderma dei tratti aortici, nel mesenchima della testa, nello sclerotoma, nel tratto di efflusso del cuore e nell'orecchio interno [62-63]. IL ruolo di *TBX1* nello sviluppo cardiovascolare è oggetto di numerosi studi. Come nei pazienti per DGS, i mutanti murini di *Tbx1* hanno due tipi di difetti cardiovascolari, nell'arco aortico e nel tratto cardiaco esterno [65-67]. Recentemente è stato dimostrato che le anomalie del tratto cardiaco esterno sono dovute ad un ruolo di *Tbx1* nell'espansione dei precursori dei cardiomiociti, che in assenza del gene, proliferano lentamente e contribuiscono in numero insufficiente al tratto cardiaco esterno. Il ridotto fenotipo proliferativo è associato alla minima espressione di *Fgf10*. Poiché è noto che *Tbx1* attiva l'espressione di *Fgf10* attraverso un elemento TBE, sito all'estremità 5' del gene stesso, è possibile concludere che *FGF10* è il primo gene noto essere bersaglio dell'attività trascrizionale di *Tbx1* [68]. Benché studi di tipo genetico abbiano prodotto molte informazioni relative al coinvolgimento di *Tbx1* nell'insorgenza della patologia, la complessità del sistema evidenzia che molti aspetti importanti devono ancora essere chiariti in merito agli effetti delle mutazioni puntiformi studiate sulla macchina cellulare. Obiettivo del progetto di ricerca è stata l'identificazione delle proteine la cui espressione dipende dall'attività del fattore trascrizionale *Tbx1* con l'ausilio di metodologie di proteomica di espressione differenziale.

Per questi studi differenziali è stata scelta, come sistema modello, la linea cellulare P19CL6, cellule embrionali di carcinoma murino. Queste cellule in vitro differenziano a cardiomiociti in seguito a trattamento con DMSO. Dopo induzione con DMSO, *Tbx1* è espresso e subisce regolazione durante il differenziamento. Un altro gene, che codifica per un fattore di trascrizione coinvolto nello sviluppo del tratto cardiaco, *Mef2c*, è espresso e regolato durante il differenziamento a cardiomiociti delle cellule P19CL6 [69]. Sono state

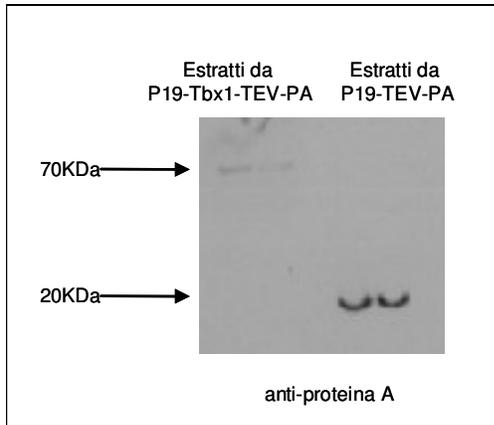
utilizzate linee cellulari, transfettate stabilmente con un plasmide esprimente Tbx1 e linee cellulari di controllo recanti solo il plasmide [70]: P19-TBX1-TEV-PA e P19-TEV-PA. Le cellule ingegnerizzate conservano la loro abilità a differenziare, ma presentano un'espressione alterata di Mef2c. Quest'ultima, infatti, è espressa già nelle cellule indifferenziate ed è regolata positivamente in seguito a trattamento con DMSO [71]. L'espressione costitutiva di Tbx1 nel sistema cellulare P19CL6 può influenzare l'espressione di geni rilevanti nel differenziamento a cardiomiociti. Dunque cellule P19CL6 esprimenti Tbx1 rappresentano un sistema modello per lo studio dei geni bersaglio dell'attività trascrizionale di Tbx1 stesso.

Per esperimenti di tipo comparativo con approccio proteomico è stata utilizzata la metodologia del DIGE. Quest'ultima consente un confronto quantitativo accurato dell'espressione proteica di cellule P19 esprimenti Tbx1 costitutivamente e cellule P19 non esprimenti Tbx1. Il protocollo sperimentale prevede la formazione di un legame covalente tra le proteine dell'estratto cellulare e coloranti fluorescenti: DyCye Cy2, Cy3 e Cy5 (Amersham). In particolare i campioni proteici da confrontare sono marcati con due dei tre fluorofori e frazionati nel medesimo esperimento di elettroforesi bidimensionale. I fluorofori hanno, infatti, spettri d'emissione e d'assorbimento non sovrapponibili, il che consente di acquisire due immagini rispettivamente per i due campioni da raffrontare, avendo tuttavia eseguito una sola corsa elettroforetica. Un campione costituito da quantità proteiche uguali dei due stati da confrontare è legato ad un terzo fluoroforo e costituisce il riferimento per la normalizzazione sperimentale (standard interno). L'intensità di fluorescenza dei due campioni sarà normalizzata rispetto allo standard interno e poi paragonata per l'analisi quantitativa. Replicati biologici dei due campioni rendono il dato statisticamente attendibile e biologicamente significativo. Essendo una tecnica di elettroforesi bidimensionale è cruciale il gradiente di pH adoperato. Per la definizione del protocollo sperimentale è stato scelto un intervallo di pH 3-11.

I risultati ottenuti dall'analisi d'immagine e dalla successiva indagine per spettrometria di massa sono stati confermati e complementati con quelli ottenuti da un secondo esperimento. Infatti, al fine di migliorare la risoluzione sperimentale è stato scelto un intervallo di pH più stretto.

## 4.2 Risultati

Per studiare le proteine la cui espressione dipende dall'attività del fattore trascrizionale Tbx1 sono stati realizzati esperimenti di proteomica comparativa su sistemi cellulari esprimenti stabilmente Tbx1 (P19-Tbx1 TEV PA) rispetto a sistemi cellulari di controllo (P19 TEV PA). Poiché il costrutto introdotto stabilmente nel sistema cellulare reca una sequenza nucleotidica codificante la proteina A è stata verificata l'espressione esclusiva di Tbx1 in cellule P19-Tbx1 TEV PA mediante western blotting anti-proteina A (figura 12).



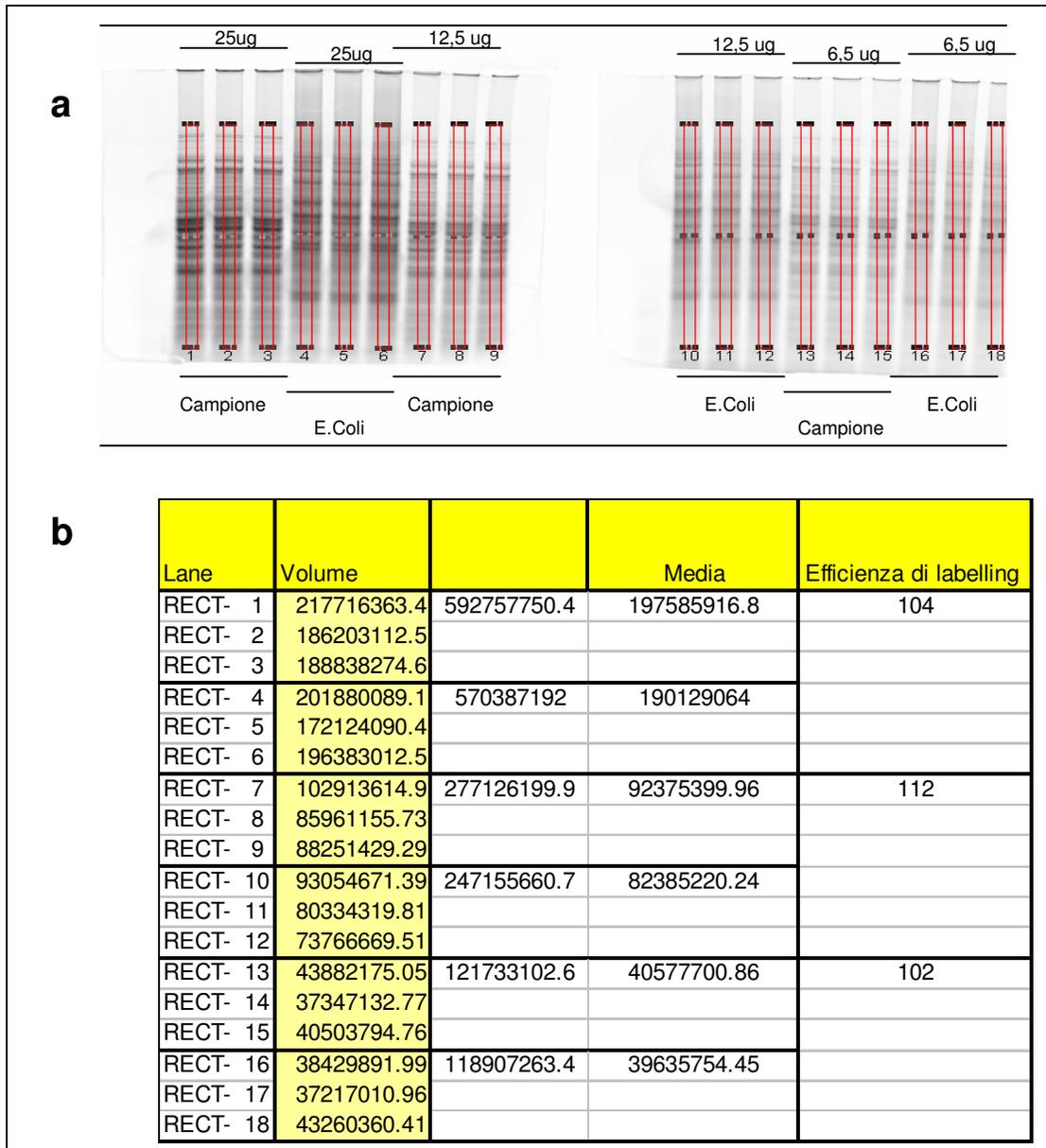
**Figura 12.** Estratti proteici di cellule P19-TBX1-TEV-PA e P19-TEV-PA sono separati su SDS-PAGE 10% e analizzati per *western blotting* con l'anticorpo secondario anti-PA.

Avvalendoci della tecnologia DIGE, è stato allestito un primo esperimento scegliendo un intervallo di pH 3-11 per la separazione elettroforetica in prima dimensione. Sono stati eseguiti due esperimenti di controllo per valutare l'efficienza di marcatura dell'estratto proteico utilizzato e per escludere la possibilità di una marcatura preferenziale del campione proteico con uno dei tre fluorofori. Nel primo esperimento il segnale di fluorescenza dell'estratto è stato confrontato con quello di un controllo, noto reagire in modo efficiente con i fluoroforo (figura 13). Nel secondo esperimento lo stesso campione è stato marcato con i tre fluorofori ed è stata eseguita l'analisi d'immagine del gel 2DE su cui la miscela di reazione è stata separata. Sono stati calcolati i rapporti tra le abbondanze relative espresse in termini di volume di fluorescenza ( $V_1/V_2$ ). Il più alto valore  $V_1/V_2$  ottenuto è stato usato come limite inferiore nell'esperimento comparativo al di sotto del quale la variazione non è ritenuta significativa (figura 14).

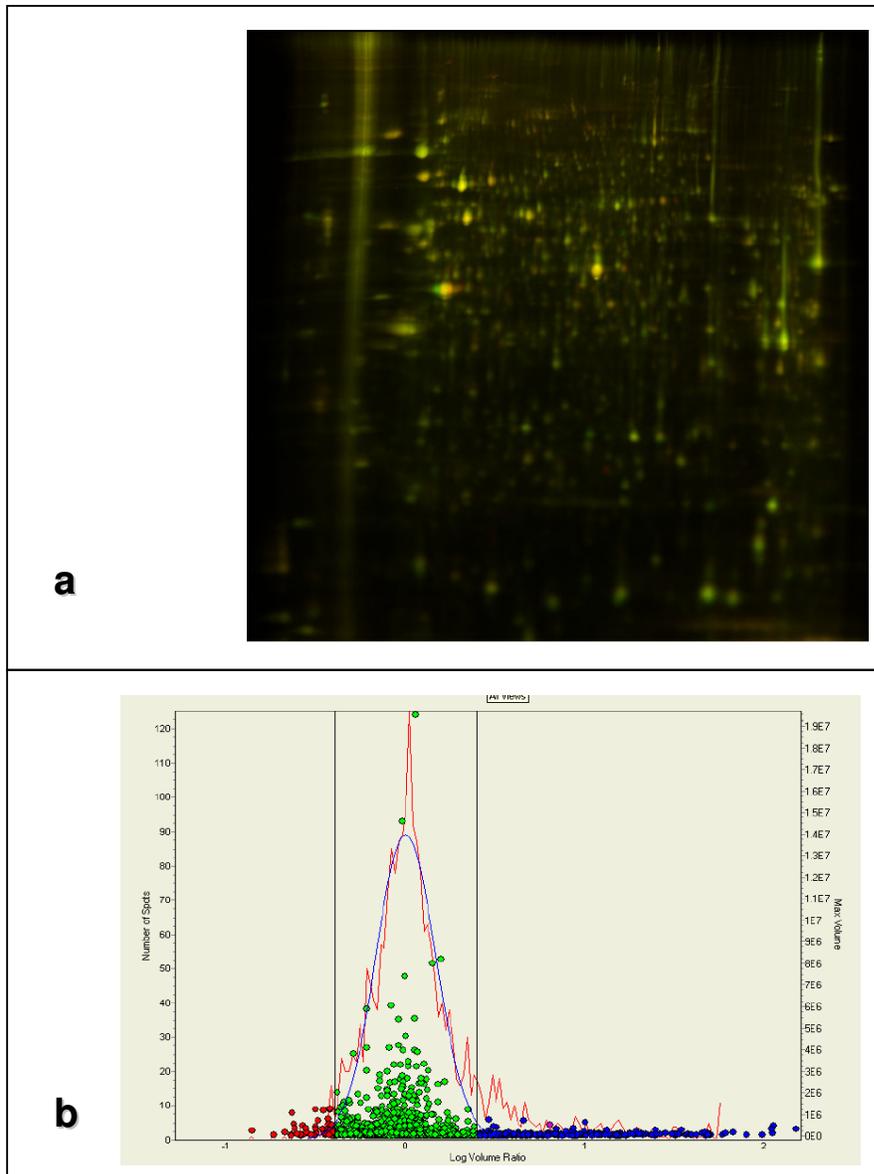
Sono stati preparati 6 replicati biologici per ognuna delle condizioni da confrontare, 6 gels analitici per l'analisi d'immagine e due gels semipreparativi per gli studi di spettrometria di massa. L'analisi d'immagine necessaria per confrontare i campioni in esame, come descritto nel paragrafo 1.3 ed in figura 6 prevede due fasi consecutive:

- ✓ co-detection intra gel (6a)
- ✓ matching inter-gel (6b).

Nella prima fase la sovrapposizione delle tre immagini acquisite in intervalli spettrali differenti sono sovrapposte. Ogni spot è intrinsecamente legato allo stesso spot presente nello standard interno. Infatti, il volume di un segnale di fluorescenza, inteso come la somma dei valori dei pixel in uno spot, sottratto del rumore di fondo è confrontato direttamente con il suo standard interno. L'abbondanza di uno spot in un campione è espressa come rapporto rispetto allo standard interno (normalizzazione).



**Figura 13.** Estratti proteici da cellule P19-TBX1-TEV-PA, campione, e da cellule di E.Coli, controllo, sono stati marcati con un fluoroforo e separati su un 1D SDS-PAGE (a). E' riportato il volume di fluorescenza registrato in ogni canale del gel, la media, e il rapporto percentuale tra i volumi (b). Quest'ultimo è indice dell'efficienza di marcatura.



**Figura 14.** Estratti proteici da cellule P19-TBX1-TEV-PA sono stati marcati con i tre fluorofori: Cy2, Cy3, Cy5 e separati su un 2D SDS-PAGE (a). E' riportato il numero di spots in funzione del Log del rapporto tra i volumi di fluorescenza ( $\text{Log } V_1/V_2$ ) del campione marcato con Cy3 e Cy5 (b).

Nella fase successiva, poiché tutti replicati contengono lo stesso standard interno, è possibile determinare il rapporto dell'abbondanza proteica tra i due campioni. La presenza dello standard interno annulla l'impatto che variazioni sperimentali hanno sull'analisi. I rapporti campione/standard interno producono un plot d'abbondanza relativa standardizzata per ogni specie proteica lungo il gruppo di replicati. In questo modo il confronto tra i due campioni è più accurato. Un test statistico (T-Student) valida il dato, assegnando un intervallo di confidenza ( $P < 0,01$ ) all'interno del quale il rapporto di abbondanze relative trovato ( $R > 1,3$  o  $R < 1,3$ ) ha significato biologico. L'analisi d'immagine condotta ha consentito di individuare diversi spots con abbondanze relative differenti nei due campioni in esame. Gli spots di interesse sono stati escissi dai corrispondenti gels semipreparativi e, dopo idrolisi in situ analizzati per spettrometria di massa tandem. Le proteine identificate sono riportate nelle tabella 5 e 6. Le identificazioni mostrano che numerosi spots di interesse constano di più d'una specie proteica. In questi casi non è stato possibile assegnare con certezza la variazione di abbondanza proteica ad una singola specie.

**Tabella 5. Identificazione delle proteine iper-esprese in cellule P19-TBX1-TEV-PA (analisi differenziale nell'intervallo di pH 3-11)**

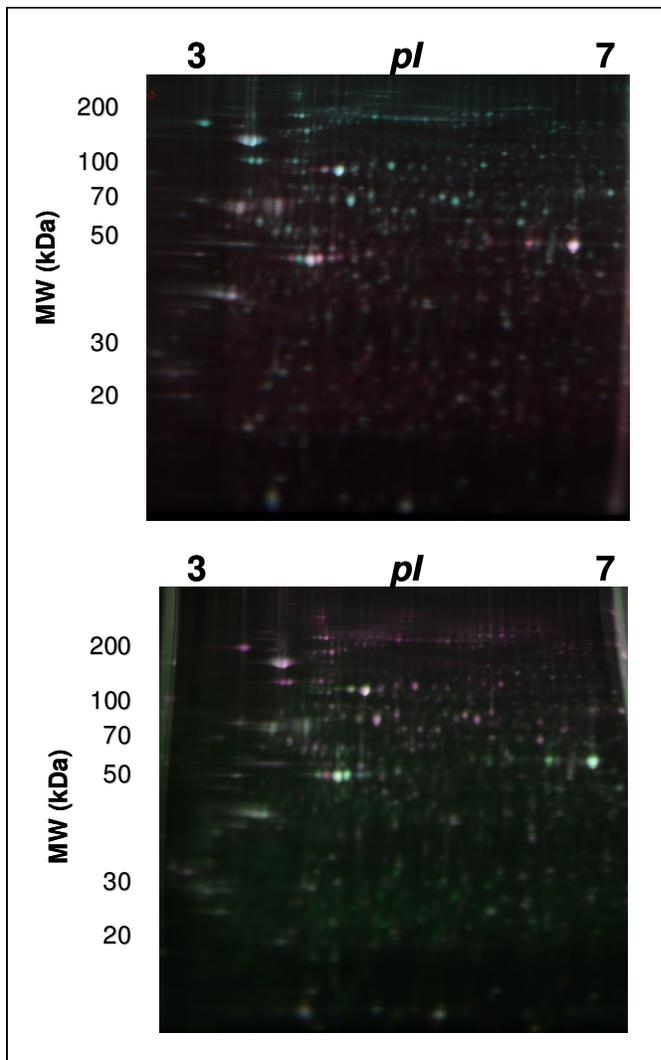
Gene	Proteina	Variazione
Nars_mouse	Asparaginyl-tRNA synthetase	0,88
Aldo1_mouse	Fructose-bisphosphate aldolase A	0,54
Hnrpd_mouse	Heterogeneous nuclear ribonucleoprotein D0	0,54
Pkm2_mouse	Pyruvate kinase, isozyme M2	0,48
CTE1_mouse	Acyl coenzyme A thioester hydrolase	0,48
Eef2_mouse	Elongation factor 2	0,44
Idi1_mouse	Isopentenyl-diphosphate delta-isomerase 1	0,39
Impdh2_mouse	Inosine-5'-monophosphate dehydrogenase 2	0,38
Prdx1_mouse	Peroxiredoxin 1	0,33
Ndufv1_mouse	NADH-ubiquinone oxidoreductase 51 kDa subunit	0,31
Shmt1_mouse	Serine hydroxymethyltransferase	0,31

**Tabella 6. Identificazione delle proteine ipo-esprese in cellule P19-TBX-TEV-PA (analisi differenziale nell'intervallo di pH 3-11)**

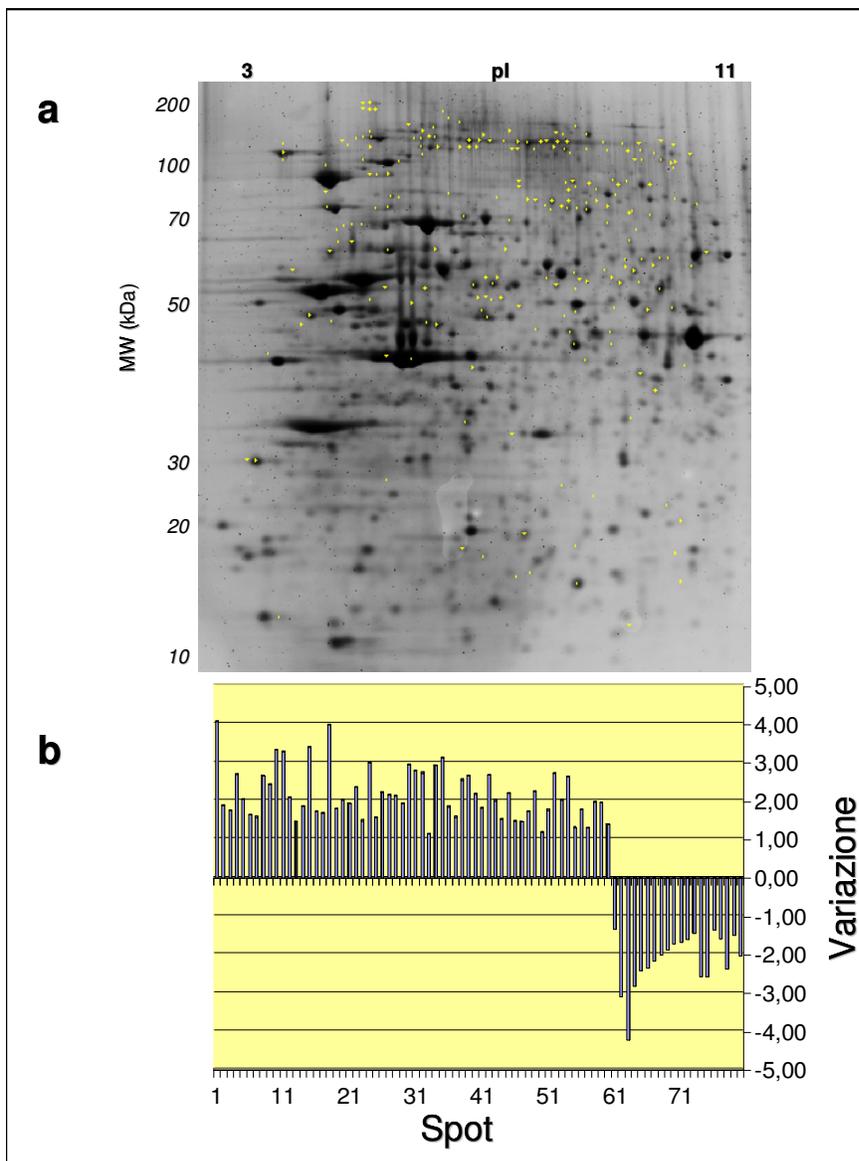
Gene	Proteina	Variazione
VDAC1_mouse	Voltage-dependent anion-selective channel protein 1	1,21
Spr_mouse	Sepiapterin reductase	0,76
Gnb1_mouse	Guanine nucleotide-binding proteinG(I)/G(S)/G(T)	0,42
Hnrph_mouse	Heterogeneous nuclear ribonucleoprotein H	0,41
Gpd2_mouse	Glycerol-3-phosphate dehydrogenase	0,31

Per migliorare la risoluzione delle mappe elettroforetiche ottenute, il medesimo esperimento è stato allestito su intervalli di pH più stretti; in particolare è stata amplificata la regione 4-7. In figura 15 e in figura 16 sono mostrati esempi di gels analitici ed uno semipreparativo eseguiti, con le variazioni trovate. La medesima procedura sperimentale realizzata su quattro replicati biologici ha condotto all'identificazione di circa 50 proteine (tabelle 7 e 8), divise in ipo o iper-esprese in conseguenza alla presenza costitutiva di Tbx1 nel sistema cellulare.

La figura 17 mostra un esempio di analisi differenziale quantitativa condotta con il programma DyCyder. Sono state identificate 37 proteine, classificate (figura 18) in gruppi funzionali (a), secondo la loro localizzazione cellulare (b), e i processi cellulari che le coinvolgono (c), utilizzando le banche dati telematiche (PubMed [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov); Mouse Genome Informatics [www.informatics.jax.org](http://www.informatics.jax.org)). La figura 19 mostra uno spettro di massa, uno spettro di frammentazione e la sequenza amminoacidica di alcune proteine identificate, PCNA e Adh2, la cui espressione proteica è alterata dalla presenza costitutiva di Tbx1 nel sistema cellulare.



**Figura 15.** Esempio di gels analitici ottenuti seguendo il protocollo sperimentale della tecnologia DIGE.



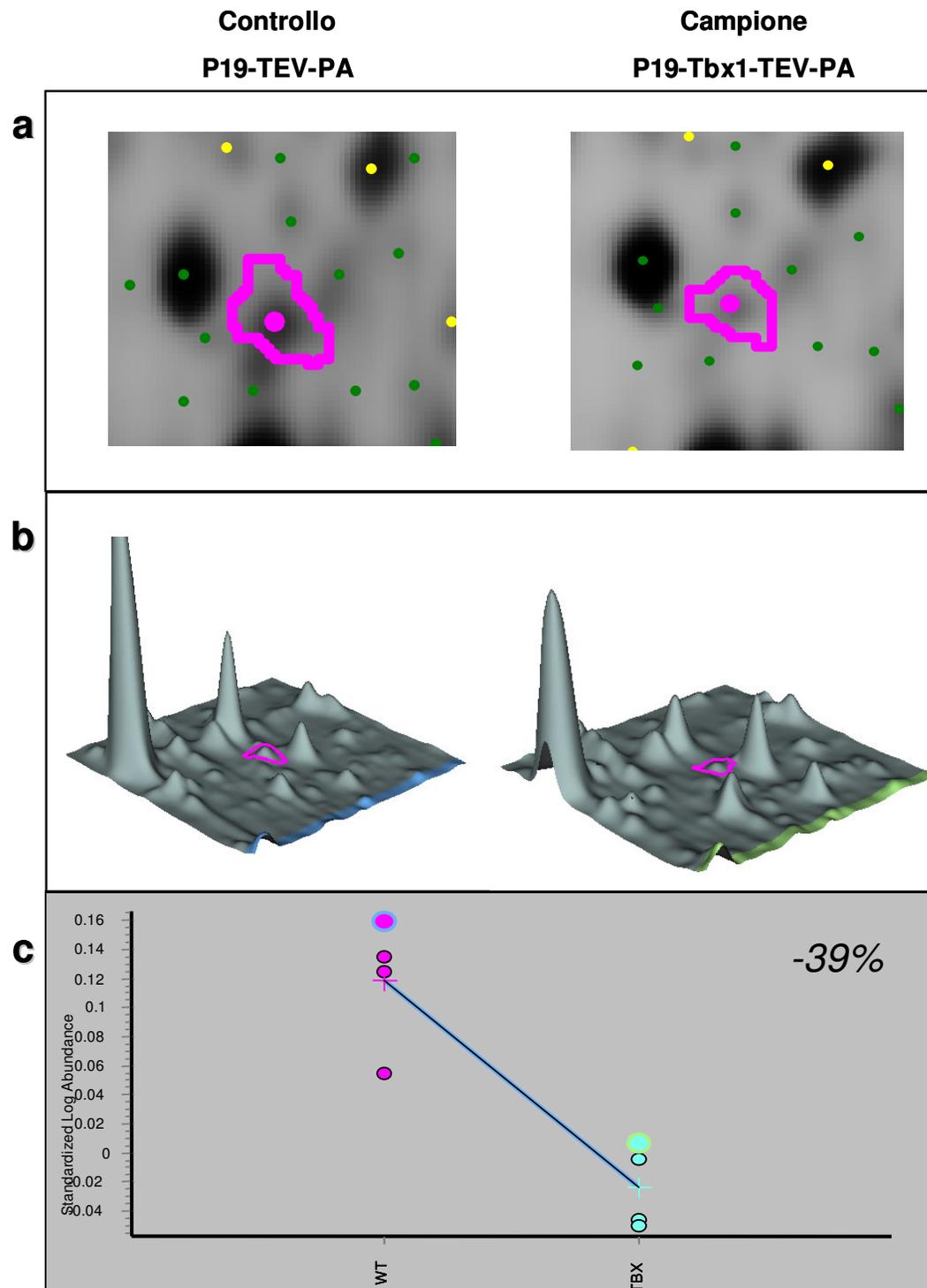
**Figura 16.** Esempio di gel analitico ottenuto seguendo il protocollo sperimentale della tecnologia DIGE (a). Analisi delle variazioni trovate nel confronto tra gli estratti di cellule P19-TBX1-TEV-PA e P19-TEV-PA (b).

**Tabella 7. Identificazione delle proteine iper-espresse in cellule P19-TBX1-TEV-PA  
(analisi differenziale nell'intervallo di pH 4-7)**

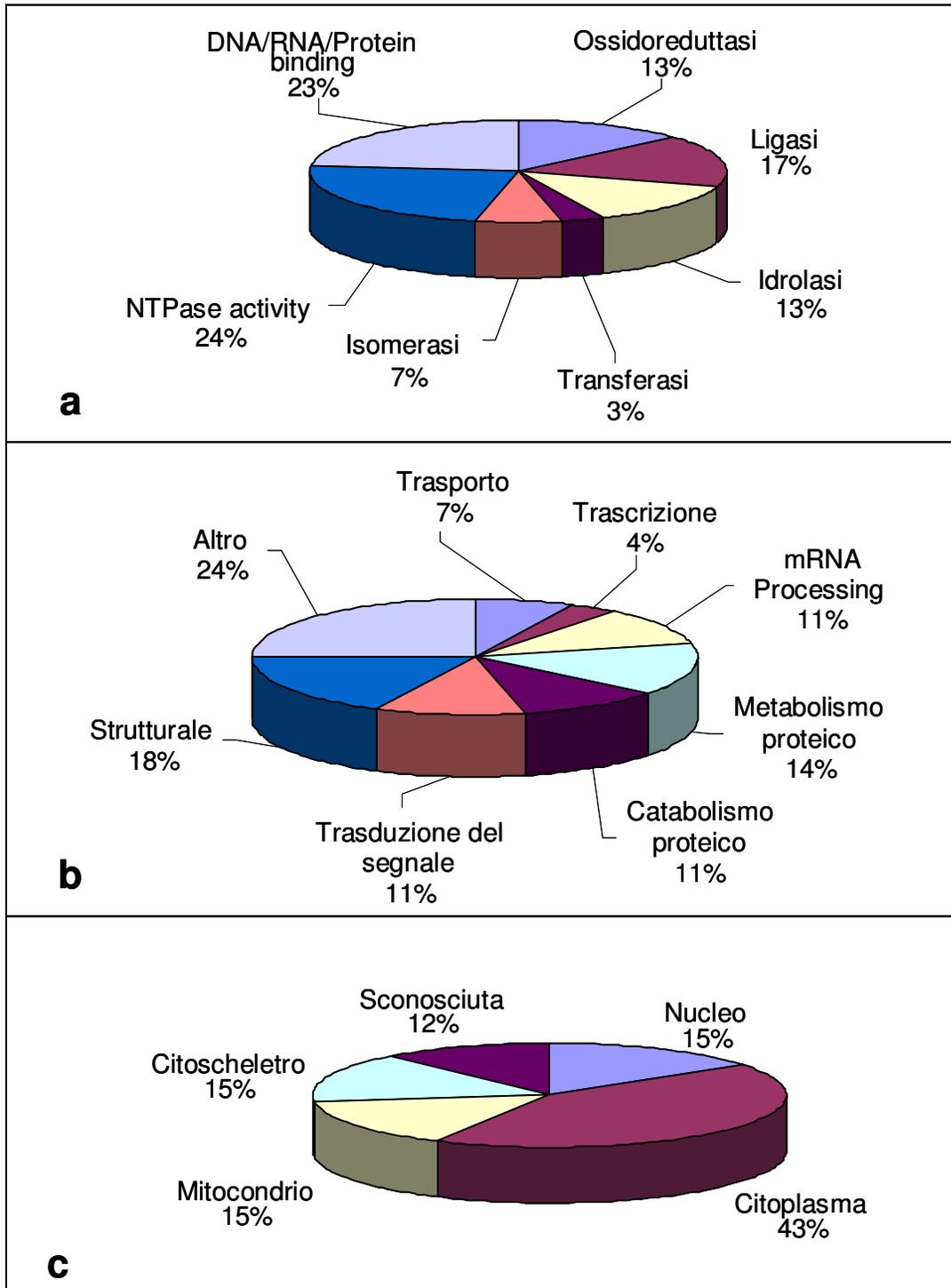
<b>Gene</b>	<b>Proteina</b>	<b>Variazione</b>
Hyou1_mouse	Hypoxia up-regulated 1	7,75
Acot2_mouse	Mitochondrial long chain acyl-CoA thioesterase	4,27
Hsp70_mouse	Heat shock 70kD protein 5	4,08
Aars_mouse	Alanyl-tRNA synthetase	3,98
Syg_mouse	Glycine-tRNA ligase	3,28
Hspa4_mouse	Heat shock protein 4	3,02
Vcl_mouse	Vinculin	2,92
Prss15_mouse	Protease, serine, 15	2,78
Ogt_mouse	UDP-N-acetylglucosaminyltransferase	2,42
Alix_mouse	Programmed cell death 6 interacting protein	2,65
Ube1X_mouse	Ubiquitin-activating enzyme E1, Chr X	2,65
Trim28_mouse	Tripartite motif protein 28	2,62
Gdi2_mouse	Guanosine diphosphate (GDP) dissociation inhibitor 2	2,46
Syncrip_mouse	Synaptotagmin-binding, cytoplasmic RNA-interacting protein	2,43
Coro7_mouse	Coronin 7	2,38
Ncl_mouse	Nucleolin	2,36
Ehd3_mouse	EH-domain containing 3	2,12
Nln_mouse	Neurolysin (metallopeptidase M3 family)	2,11
Ganab_mouse	Alpha glucosidase 2 alpha neutral subunit	2,09
Sys_mouse	Serine-tRNA ligase	2,04
Fkbp4_mouse	FK506 binding protein 4	1,64
PBP74_mouse	Similar to Stress-70 protein, mitochondrial precursor	1,89
Vil2_mouse	Villin 2	1,78
Immt_mouse	Inner membrane protein, mitochondrial	1,78
Asns_mouse	Asparagine synthetase	1,73
Phgdh_mouse	3-phosphoglycerate dehydrogenase	1,68
Pdia3_mouse	Protein disulfide isomerase associated 3	1,58
Dync1li1_mouse	Dynein, cytoplasmic, light intermediate chain 1	1,56
Ahcy_mouse	S-adenosylhomocysteine hydrolase	1,46
Pgm2_mouse	Phosphoglucomutase 2	1,43

**Tabella 8. Identificazione delle proteine ipo-esprese in cellule P19-TBX-TEV-PA  
(analisi differenziale nell'intervallo di pH 4-7)**

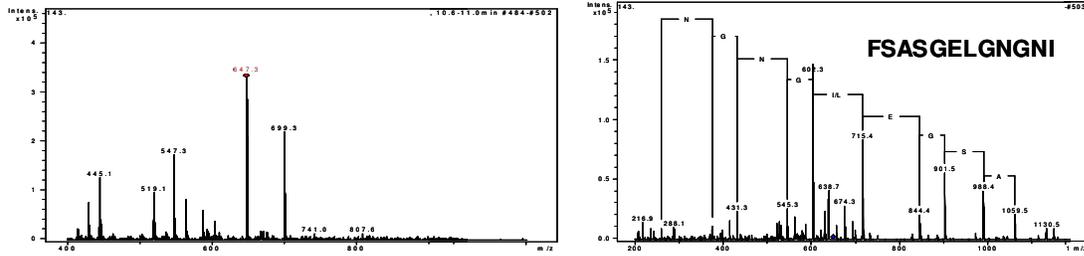
<b>Gene</b>	<b>Proteina</b>	<b>Variazione</b>
Hmgcs1_mouse	3-hydroxy-3-methylglutaryl-Coenzyme A synthase 1	-2,82
Adprh_mouse	ADP-ribosylarginine hydrolase	-2,31
Psme1_mouse	Proteasome (prosome, macropain) 28 subunit, alpha	-2,18
Pcna_mouse	Proliferating cell nuclear antigen	-1,67
Oat_mouse	Oat protein (ornithine aminotransferase)	-1,61
Sept02_mouse	Septin 2	-1,53
Aldh2_mouse	Aldehyde dehydrogenase 2, mitochondrial	-1,39



**Figura 17.** Espressione della proteina Adh2 nei due campioni analizzati: estratti da cellule P19-TEV-PA (controllo) e cellule P19-TBX1-TEV-PA, esprimenti costitutivamente Tbx1 (TBX). Visualizzazione della variazione su gel 2D (a). Rappresentazione tridimensionale del segnale di fluorescenza relativo allo spot (b). Abbondanza proteica relativa in funzione dei campioni studiati (c).

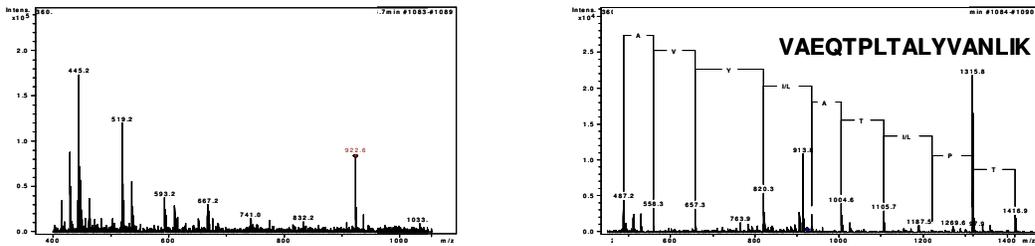


**Figura 18.** Le proteine, la cui espressione proteica varia in cellule esprimenti costitutivamente Tbx1 rispetto al controllo, sono state classificate secondo la loro funzione molecolare (a), localizzazione cellulare (b), processo biologico (c), in accordo con la banca dati Gene Ontology (PubMed [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)).



**a**

MFEARLIQGS ILKKVLEALK DLINEACWDV SSGVNLQSM DSSHVSLVQL TLRSEGFDR CDRNLAMGVN  
 LTSMSKILKC AGNEDIITLR AEDNADTLAL VFEAPNQEKV SDYEMKMLDL DVEQLGIPEQ EYSCVIKMP  
 GEFARICRDL SHIGDAVVIS CAKNGVKFSA SGELGNGNIL SQTSNVDKKEE EAVTIEMNEP VHLTFALRYL  
 NFFTKATPLS PVTLSMSAD VPLVVEYKIA DMGHLKYLLA PKIEDEEAS



**b**

MLRAALTIVR RGPRLSRLLS AAATSAVPAP NHQPEVFCNQ IFINNEWHDA VSRKTFPTVN PSTGEVICQV  
 AEGNKEDVDK AVKAARAAPQ LGSPWRRMDA SDRGRLLYRL ADLIERDRTY LAALLETLDNG KPYVISYLVLD  
 LDMVLKCLRY YAGWADKYHG KTIPIIDGDF SYTRHEPVGV CGIIPWNFP LLMQAWKLG ALATGNVVVM  
KVAEQTPLTA LYVANLIKEA GFPPGVVNI V PGFGPTAGAA IASHEGVDKV AFTGSTEVGH LIQVAGSSN  
 LKRVTLLEGG KSPNIIMSDA DMDWAVEQAH FALFFNQGQC CCAGSRFTVQ ENVYDEFVER SVARAKSRV  
 GNPFDSTRTEQ GPQVDETQFK KILGYIKSGQ QEGAKLLCGG GAAADRGYFI QPTVFGDVKD GMTIAKEEIF  
 GPVMQILKFK TIEEVVGRAN DSKYGLAAAV FTKDLDKANY LSQALQAGTV WINCYDVFGA QSPFGGYKMS  
 GSGRELGEYG LQAYTEVKTV TVKVPQKNS

**Figura 19.** Spettro di massa, spettro di frammentazione e sequenza amminoacidica identificata per le proteine PCNA (a), Adh2 (b).

### 4.3 Discussione

Tbx1 è un fattore trascrizione coinvolto nella sindrome di DiGeorge: mutazioni puntiformi nella sua sequenza genica sono state ritrovate in pazienti affetti dalla patologia. Pur essendo chiaro il coinvolgimento di Tbx1 nell'insorgenza della patologia, non sono ancora noti gli effetti delle mutazioni puntiformi studiate sulla macchina cellulare. E' stato condotto uno studio di proteomica differenziale al fine di individuare i bersagli dell'attività trascrizionale di Tbx1, in cellule embrionali di carcinoma murino P19, esprimenti la proteina in un numero di copie superiore rispetto al livello fisiologico. E' stata scelta la tecnologia DIGE per individuare e quantizzare variazioni nell'espressione proteica dei campioni analizzati. E' stato condotto un primo esperimento scegliendo un gradiente di pH 3-11; quest'ultimo ha consentito la definizione dei parametri del protocollo sperimentale e ha condotto all'identificazione di 16 specie proteiche la cui espressione si manifestava variata nei campioni confrontati. Al fine di migliorare la risoluzione sperimentale e minimizzare fenomeni di co-migrazione elettroforetica e assegnare in maniera univoca le variazioni proteiche trovate è stato allestito un secondo esperimento su gradienti di pH più stretti. In particolare è stato scelto l'intervallo 4-7. L'analisi d'immagine e per spettrometria di massa condotta, ha consentito di individuare ed identificare 37 specie proteiche, i cui livelli di espressione mostrano abbondanze relative differenti nei due campioni in esame. Le proteine identificate nei due esperimenti condotti sono state classificate (figura 18) secondo la loro specificità funzionale (a), la loro localizzazione cellulare (b), e i processi cellulari che le coinvolgono (c).

Le funzioni biologiche assegnate mostrano che molte delle proteine identificate sono coinvolte in processi di nucleotide tri-fosfatasi (NTPasi) e di legame a DNA, RNA e proteine. L'attività fosfataseica induce ad ipotizzare un coinvolgimento delle proteine identificate nella regolazione dei processi di trasduzione del segnale, ipotesi avvalorata dall'attività di legame ad altre specie proteiche. Il legame a DNA e RNA suggerisce un legame con processi di trascrizione e splicing.

Molte proteine (b) sono coinvolte in processi di trasporto e in ruoli strutturali, altre in processi di trascrizione, trasduzione e di processing dell'mRNA. La maggior parte ha una localizzazione citoplasmatica; altre sono equamente divise tra nucleo, mitocondrio e citoscheletro.

Tra le proteine ipoespresse nei lisati proteici di cellule over-esprimenti Tbx1, l'*aldehyde dehydrogenase 2* (Adh2) ha destato particolare interesse, essendo un enzima coinvolto nell'ossidazione della retinaldeide ad acido retinoico [72], la cui concentrazione molare è critica per lo sviluppo dell'apparato faringeo. I suoi livelli d'espressione diminuiscono, come riportato in figura 17, per il 39%, indicando una ipo-regolazione diretta o indiretta da parte di Tbx1. La sua concentrazione cellulare è il risultato dell'equilibrio tra attività anaboliche e cataboliche. Lo spostamento dell'equilibrio verso una delle due attività porta a malformazioni embrionali [73]. E' stato dimostrato che la mancanza di Tbx1 porta ad un aumento della produzione di acido retinoico [74]. Poiché l'aumento di Tbx1 provoca una ipo-espressione dell'*aldehyde dehydrogenase 2* si potrebbe ipotizzare un meccanismo di regolazione mediato da Tbx1, finalizzato al controllo della concentrazione di acido retinoico.

Una seconda proteina d'interesse è la *proliferating cell nuclear antigen* (PCNA), ipoespressa del 67% in condizioni di eccesso di Tbx1. E' una ciclina presente ad alte concentrazioni in cellule in divisione. E' noto dalla letteratura che, in embrioni, i livelli di trascritto codificante la proteina aumentano in risposta ad una stimolazione con acido retinoico [75]. La relazione tra la concentrazione di acido retinoico e PCNA è stata, tuttavia, valutata soltanto nei livelli di trascritto. Inoltre è stato evidenziato che l'interattore

di PCNA con il recettore dell'acido retinoico (RAR), ne reprime la funzione [76]. Probabilmente, nel sistema sperimentale preso in esame, l'abbassamento dei livelli di acido retinico provocato dalla regolazione negativa di Tbx1 sull'aldehyde dehydrogenase 2 potrebbe portare ad una diminuzione dei livelli del recettore RAR e alla conseguente diminuzione dei livelli di PCNA.

I risultati confermano l'elevata affidabilità della tecnologia DIGE come metodologia per studi di proteomica differenziale, e rappresentano un esempio di approccio di proteomica differenziale per lo studio dell'azione di un fattore di trascrizione. L'analisi proteomica differenziale è un'ottima strategia, che fornendo una visione d'insieme del sistema preso in esame offre spunti importanti per successivi studi di approfondimento. I risultati ottenuti necessitano di essere validati e complementati con tecniche di biochimica classica e avanzata, quali *western blotting* per la conferma dell'espressione proteica differenziale, saggi di *chromatin-immuno-precipitation* e analisi *microarray* dei trascritti cellulari per la validazione dei bersagli dell'attività trascrizionale di Tbx1.

## Capitolo 5

# ANALISI DELL'INTERATTOMA DELLA PROTEINA RIBOSOMIALE S19

### 5.1 Introduzione

Rps19 è una proteina strutturale di 17KDa appartenente alla subunità 40S del ribosoma. Mutazioni puntiformi nella sua sequenza nucleotidica sono state identificate in soggetti affetti da una rara malattia ematologica, la Diamond Blackfan Anemia (DBA) (MIM 105650) [77-79]. Il 25% dei soggetti patologici per DBA presenta mutazioni nel gene RPS19. La DBA è una malattia congenita, caratterizzata da una maturazione imperfetta dei progenitori eritroidi ed è la prima patologia umana dovuta a mutazioni in una proteina ribosomiale essenzialmente strutturale. Non esiste nessuna relazione tra manifestazione clinica, risposta al trattamento e genotipo. Il ruolo di Rps19 nell'eritropoiesi fisiologica e l'impatto sull'anomala eritropoiesi in DBA non è ancora stato definito [80], ma l'osservata soppressione dell'espressione di un allele in molte delle mutazioni studiate suggerisce che la principale causa del fenotipo patologico è l'aploinsufficienza del gene stesso.

Come altre proteine ribosomiali, Rps19 trasloca dal citoplasma al nucleo, ed in particolare nel nucleolo. Non sono state definite sequenze amminoacidiche che determinano la traslocazione nucleolare, sembra piuttosto che i determinanti molecolari di suddetta traslocazione siano la struttura terziaria delle proteine e l'interazione tra i loro domini funzionali ed altre proteine nucleolari o molecole di rRNA [81-83]. Nel nucleo Rps19 partecipa alla biogenesi del ribosoma. Recenti studi hanno dimostrato che l'assenza in lievito di Rps19 determina un'alterata maturazione degli rRNA e una sintesi anomala della subunità ribosomiale 40S [84-85].

E' noto infine che Rps19 interagisce con PIM, una serin treonin chinasi ubiquitaria, la cui espressione può essere indotta in cellule eritropoietiche da fattori di crescita. PIM interagisce con il ribosoma e potrebbe essere coinvolta in processi di controllo traduzionale [86]. Suddetta interazione suggerisce un ruolo in processi di controllo traduzionale anche per Rps19, come mostrato per altre proteine ribosomiali: Rpl13 e Rpl26 [87-88]. Dunque Rps19 potrebbe svolgere, oltre a quella puramente strutturale, nuove funzioni, molte delle quali in associazione con complessi multiproteici. La sua funzione e regolazione può essere chiarita dalla conoscenza dei suoi ligandi proteici mediante strategie sperimentali di proteomica funzionale. Il progetto è stato svolto in collaborazione con il gruppo di ricerca della prof. Irma Dianzani (Università di Novara).

#### 5.1.2 Biogenesi del ribosoma

Negli eucarioti la sintesi del ribosoma avviene nel nucleolo, compartimento nucleare. Il processo comincia con la sintesi dei pre rRNA 5S e 35S, che coinvolgono RNA polimerasi distinte e richiedono l'importo delle proteine ribosomiali dal citoplasma. Il rilascio delle componenti di RNA mature dal pre rRNA richiede il coinvolgimento di idrolisi endo ed eso nucleolitiche. Inoltre i pre rRNA sono variamente modificati e legati a proteine ribosomiali prima che le subunità pre-40S e pre-60S assemblate siano esportate al citoplasma [89]. Le varie proteine ribosomiali e gli rRNA sono prodotti in quantità equimolari e la loro sintesi è fortemente regolata da una varietà di fattori di crescita [90]. La maturazione degli rRNA

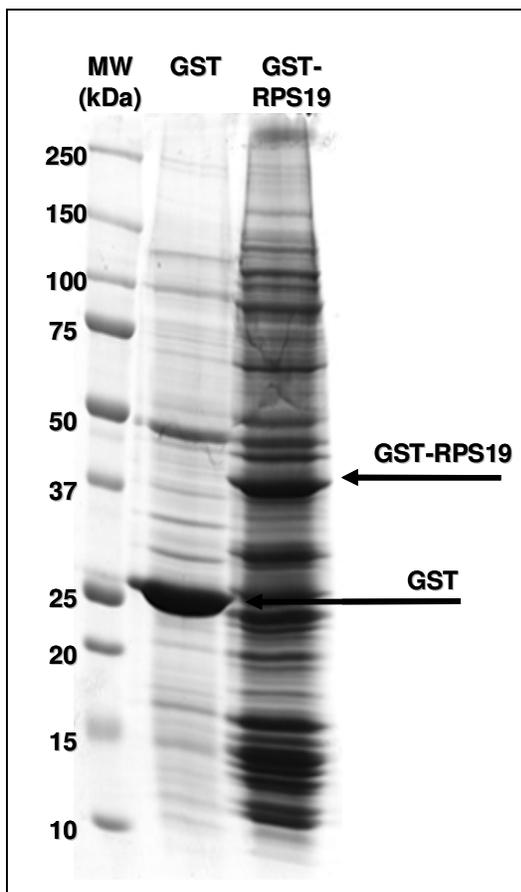
ribosomiali e il loro assemblaggio nelle subunità del ribosoma coinvolgono 170 proteine accessorie tra cui *endo ed esonuclease*, *ATP-dependent RNA-elicase*, *chaperone* e fattori di assemblaggio [89; 91] e molte *small nucleolar ribonucleoprotein* (snoRNPs). L'ampio numero è collegato al peculiare ed elaborato sistema di metilazione del ribosio e formazione di pseudouridine negli eucarioti, in cui la specificità dei siti bersaglio è collegata all'ibridazione dell'RNA con snoRNPs [92-93].

La formazione dei ribosomi negli eucarioti non solo richiede la coordinazione di eventi di assemblaggio e maturazione, ma anche un ordine spazio temporale che definisce un inizio nel nucleolo ed una fine nel citoplasma [94]. Molte altre proteine sono addette alla coordinazione tra i processi di assemblaggio e l'esporto dal nucleo. Con saggi di ricostituzione *in vitro* del sistema ribosomiale batterico si evince che il processo di assemblaggio delle subunità ribosomiali a partire dalle componenti, proteine individuali e molecole di rRNA, non richiede la presenza di cofattori, ma solo uno step di riscaldamento che abbassa l'energia di attivazione del processo. *In vivo* l'assemblaggio richiede la partecipazione di pochi cofattori, tra cui *ATP-dependent RNA-elicase* [95] e dello chaperone DnaK/hsp70 [96-97]. Negli eucarioti il medesimo sistema è molto più complicato e richiede nei processi di maturazione, assemblaggio ed esporto dal nucleo il coinvolgimento di molti fattori. Solo recentemente l'uso di approcci proteomici [6,98-99] ha consentito, attraverso l'uso di tecniche di purificazione con resine derivatizzate da componenti proteici coinvolti nella biogenesi ribosomiale, l'identificazione di una serie di proteine e la definizione di una rete di interazioni che aprono nuove prospettive nel problema della correlazione tra composizione proteica del nascente pre-ribosoma e un dato stato di maturazione degli rRNA [100-102].

## 5.2 Risultati

### 5.2.1 Identificazione dei ligandi proteici di Rps19

Per determinare l'interattoma della proteina Rps19 in cellule umane K562, è stata eseguita un'analisi LCMS/MS delle proteine purificate da un esperimento di cromatografia d'affinità su resine di sefarosio derivatizzate con GST-Rps19 e GST. Le proteine eluite dall'esca GST-Rps19 e dal controllo negativo, l'esca GST, sono state frazionate su SDS PAGE in gradiente 8-18% e rivelate mediante colorazione con Coomassie Colloidale (figura 20). Il gel mostra che le proteine associate a Rps19 coprono un ampio intervallo di peso molecolare. E' possibile visualizzare due principali bande, rispettivamente a 26KDa e 43 KDa che corrispondono alle proteine esca utilizzate nelle due purificazioni. Sono state ottenute 65 miscele peptidiche per ogni purificazione. Le miscele peptidiche che derivano dalla purificazione con GST rappresentano il controllo negativo per l'analisi delle miscele ottenute da GST-Rps19.



**Figura 20.** Estratti proteici da cellule K562 sono purificati mediante cromatografia d'affinità con resine derivatizzate con GST e GST-RPS19. Gli eluati sono stati frazionati su gel d'acrilammide 8-18% e rivelati mediante colorazione con Coomassie Colloidale.

Per selezionare le proteine che interagiscono in modo specifico con Rps19 sono state sottratte le specie comuni alle due purificazioni. Le proteine identificate solo dall'analisi del canale relativo alla purificazione con GST-Rps19 sono state raccolte nella tabella 9. Sono state raggruppate in base alla loro funzione e per ogni identificazione è associato il gene corrispondente ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)) e il gene ortologo in *S.Cerevisiae* ([www.ensembl.org](http://www.ensembl.org)). La tabella 10 elenca le proteine presenti anche nella purificazione di

controllo, condotta con l'esca GST. Le tabelle 11 e 12 mostrano per ogni specie proteica identificata, rispettivamente nel campione e nel controllo, il numero di peptidi utilizzato, e per ogni peptide, la sequenza, l'errore sulla massa e lo *Score Mascot*. Proteine identificate da una sola sequenza peptidica sono state controllate con il programma di ricerca BLAST (NCBI; <http://ncbi.nlm.nih.gov/blast>) contro la classe tassonomica umana. Le proteine, la cui ricerca eseguita con i soli parametri della sequenza amminoacidica e della classe tassonomica restituiva una identificazione non ambigua, sono state inserite nella lista degli interattori. La figura 21 mostra uno spettro MS, uno spettro MS/MS e la sequenza amminoacidica di tre proteine, identificate tra gli interattori di Rps19: IGF2BP1 (21a), MCM6 (21b), DDX5 (21c). Infine nella figura 22 sono mostrati esempi di proteine identificate da un solo peptide: CCT2 (22a), DDX17 (22b), NOLA3 (22c). Sono state identificate 159 proteine, classificate (figura 23) in gruppi funzionali (23a), secondo la loro localizzazione cellulare (23b), e i processi cellulari che le coinvolgono (23c). La suddivisione funzionale mostra la presenza di: ATP/GTPasi (5 proteine), idrolisi/elicasasi (19 proteine), isomerasi (2 proteine), chinasi (3 proteine), fattori di splicing (5 proteine), costituenti strutturali del ribosoma (29 proteine), fattori trascrizionali (11 proteine), transferasi (5 proteine), trasportatori (9 proteine), proteiche che legano DNA/RNA (53 proteine), ed altre tra cui deidrogenasi (1 proteina), ligasi (1 proteina), peptidasi (1 proteina), recettori (1 proteina), fattori di elongazione (1 proteina) e 13 proteine di funzione non nota.

Tabella 9. Classificazione funzionale dei ligandi proteici di Rps19

Gene	Proteina	Gene ortologo in <i>S.Cerevisiae</i>
<b>NTPase activity</b>		
GTPBP4	GTP/binding protein NGB (G protein binding CRFG)	NOG1
MCM6	p105MCM (MCM6 minichromosome maintenance deficient 6)	MCM6
MCM7	p85MCM protein (MCM7 minichromosome maintenance deficient 7)	CDC47
PSMC5	proteasome 26S ATPase subunit 5	RPT6
PSMC6	proteasome 26S ATPase subunit 6	RPT4
RAB11B	RAB11B member RAS oncogene family	YPT31
XAB1	XPA binding protein 1, GTPase	NPA3
<b>Hydrolase/helicase activity</b>		
DDX5	growth regulated nuclear 68 protein (DEAD box polypeptide 5)	DBP2
DDX17	DDX17 protein	
DDX18	RNA helicase (DEAD box polypeptide 18)	HAS1
DDX21	RNA helicase II / Gu protein (DEAD box polypeptide 21)	
DDX24	DEAD box polypeptide 24	MAK5
DDX3X	dead box , X isoform (DEAD box polypeptide 3)	
DDX41	DEAD box protein astrakt	
DDX50	DEAD box polypeptide 50 (Nucleolar protein GU2)	
DDX54	ATP/dependent RNA helicase (DEAD box polypeptide 54)	DBP10
DHX9	RNA helicase A (DEAH box polypep. 9)	
DHX15	DEAH box polypeptide 15	PRP43
<i>DHX30</i>	<i>DEAH (Asp-Glu-Ala-His) box</i>	
DHX36	DEAH box polypeptide 36	
PPP2R1A	PPP2R1A (Ser/Thr protein phosphatase 2A)	
RUVBL2	RuvB-like 2	RVB2
SKIV2L2	superkiller viralicidic activity 2-like 2 (S. cerevisiae)	MTR4
SMARCA5	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 5	ISW2
XRN2	Dhm1-like protein (5'-3' exoribonuclease 2)	RAT1
<b>Isomerase activity</b>		
DKC1	Cbf5p homolog (dyskerin)	
PPIH	peptidyl prolyl isomerase H	
<b>Kinase activity</b>		
CSNK2A1	casein kinase 2, alpha 1 polypeptide	
SRP72	signal recognition particle 72	SRP72
PRKCQ	delta 2 isopentenyl pyrophosphate transferase like protein (protein kinase C, theta?)	

<b>Splicing factor activity</b>		
FUSIP1	FUS interacting protein (serine/arginine rich) 1	
SF3B1	splicing factor 3b, subunit 1, 155kDa	HSH155
SF3B3	splicing factor 3b, subunit 3, 130kDa	RSE1
SFRS10	splicing factor arg/ser rich 10	

<b>Structural constituent of Ribosome</b>		
RPL10A	60S ribosomal protein L10a	RPL1B
RPL14	60S ribosomal protein L14	RPL4B
RPL24	60S ribosomal protein L24	RPL24A
RPL27A	60S ribosomal protein L27a	RPL28
RPL3	60S ribosomal protein L3	RPL3
RPL4	60S ribosomal protein L4	RPL4B
RPL6	60S ribosomal protein L6	RPL6B
RPL7	60S ribosomal protein L7	
RPL7A	60S ribosomal protein L7a	RPL8B
RPL8	60S ribosomal protein L8	RPL2A
RPL9	60S ribosomal protein L9	RPL9B
RPLP0	60S ribosomal protein P0	RPP0
RPLP1	60S acidic ribosomal protein P1 isoform 1	RPP1A
RPLP2	60S acidic ribosomal protein P2	RPP2B
RPS10	40S ribosomal protein S10	RPS10A
RPS14	40S ribosomal protein S14	RPS14B
<i>RPS1</i>	<i>40S ribosomal protein S16</i>	<i>RPS16A</i>
RPS2	40S ribosomal protein S2	RPS2
RPS23	40S ribosomal protein S23	RPS23A
RPS24	40S ribosomal protein S24	RPS24B
RPS26	40S ribosomal protein S26	RPS26B
RPS3	40S ribosomal protein S3	RPS3
RPS4X	40S ribosomal protein S4, X-linked	RPS4A
RPS5	40S ribosomal protein S5	RPS5
RPS6	40S ribosomal protein S6	RPS6B
RPS7	40S ribosomal protein S7	RPS7A
RPS8	40S ribosomal protein S8	RPS8B
RPSA	ribosomal protein SA	RPS0A
RSL1D1	PBK1 protein	

<b>Transcription factor</b>		
BAZ1B	bromodomain adjacent to zinc finger domain, 1B	
HNRPD	heterogeneous nuclear ribonucleoprotein D2	
ILF2	interleukin enhancer binding factor 2	
ILF3	nuclear factor associated with dsRNA NFAR-2	
NKRF	transcription factor NRF	
PURA	purine-rich element binding protein A	
TAF15	TLS protein (TBP-associated factor 15)	NPL3
TRIM28	tripartite motif-containing 28	
UBTF	upstream binding transcription factor, RNA polymerase I	
XPO5	exportin 5	MSN5
YBX1	DNA binding protein B	

<b>Transferase activity</b>		
FDFT1	farnesyl-diphosphate farnesyltransferase	ERG9
FTSJ3	FtsJ homolog 3 (E. coli)	SPB1
NAT10	N-acetyltransferase 10 (FLJ10774)	
NOL1	proliferating cell nuclear protein p120	NOP2

ZC3HAV1 (NOL protein 1)  
Zinc finger CCCH type, antiviral 1

<b>Transporter activity</b>		
COPA	coatamer protein complex, subunit alpha	COP1
COPB2	coatamer protein complex, subunit beta 2	SEC27
CSE1L	CSE1 chromosome segregation 1-like (yeast)	CSE1
IPO4	importin 4	KAP123
IPO7	importin 7	
NPEPL1	aminopeptidase like 1	
STAU1	Staufen protein	
SSR4	signal sequence receptor delta	
XPO1	exportin 1 (CRM1 homolog yeast)	CRM1
<b>DNA/RNA/protein binding capacity</b>		
AATF	Ded protein (Apoptosis antagonizing transcription factor )	BFR2
ACTR1B	ARP1 actin-related protein 1 homolog B, centractin beta	
C1orf77	DKFZP547E1010 protein	
CCT2	chaperonin containing TCP1, subunit 2 (beta)	CCT2
CCT8	chaperonin containing TCP1, subunit 8 (theta)	CCT8
CEBPZ	CCAAT/enhancer binding protein zeta	MAK21
CENPC1	centromere protein C 1	
COPG	coatamer protein complex, subunit gamma 1	SEC21
DNAJC9	DnaJ homolog, subfamily C, member 9	
FBL	fibrillarin, U3 small nucleolar interacting protein 1	NOP1
GNB2L1	Guanine nucleotide binding protein (G protein), beta polypeptide 2-like 1	
HIST1H1C	histone H1b	
HIST1H1D	Histone H1 member 3	
HIST1H2AK	H2A histone family	
HIST1H2BL	H2B histone family, member C	
HIST1H2BO	histone 1, H2bo (histone H2A.2 ?)	
HIST2H4	histone 2 H4	
HNRPA2B1	heterogeneous nuclear ribonucleoprotein A2/B1	
HNRPA3	heterogeneous nuclear ribonucleoprotein A3	
HNRPC	heterogeneous nuclear ribonucleoprotein C	
HNRPD	heterogeneous nuclear ribonucleoprotein D-like (A+U-rich element RNA binding factor)	
HNRPF	heterogeneous nuclear ribonucleoprotein F	
HNRPR	heterogeneous nuclear ribonucleoprotein R	
HNRPU	heterogeneous nuclear ribonucleoprotein U (scaffold attachment factor A)	
HNRPUL2	heterogeneous nuclear ribonucleoprotein U-like 2	
HP1BP3	HP1-BP74	
IGF2BP1	IGF-II mRNA-binding protein 1	
IGF2BP3	Koc1 (IGF-II mRNA-binding protein 3)	
IMP3	U3 snoRNP protein 3 homolog	IMP3
LYAR	hypothetical protein FLJ20425	YCR087C-A
NCL	Nucleolin	NSR1
NIP7	60S ribosome subunit biogenesis protein Nip7 homolog (S. cerevisiae)	NIP7
NOLA1	nucleolar protein family A member 1	GAR1

NOL5A	(H/ACA small nucleolar RNPs) hNop56	SIK1
PABPC3	poly(A) binding protein	
PAK1IP1	PAK/PLC-interacting protein 1	MAK11
PNN	pinin, desmosome associated protein	
RAB1B	RAB1B member RAS oncogene family	
RAP1B	RAP1B member RAS oncogene family	RSR1
RBM19	RNA binding motif protein, Y-linked, family1,1 member A1	
RBMX	RNA binding motif protein, X-linked (heterogeneous nuclear ribonucleoprotein G)	MRD1
RNPC2	RNA-binding region containing protein 2	
SART3	squamous cell carcinoma antigen recognised by T cells 3	
SF3B2	splicing factor 3b, subunit 2, 145kDa	CUS1
SFRS9	splicing factor Arg/Ser 9	
SNRPA1	small nuclear ribonucleoprotein polypeptide A' (U2 small nuclear ribonucleoprotein polypeptide A')	
SNRPG	small nuclear ribonucleoprotein polypeptide G	SMX2
SNRPN	small nuclear ribonucleoprotein polypeptide N	
SRP68	signal recognition particle 68	SRP68
SURF6	surfeit protein 6	RRP14
SYNCRIP	NS1 associated protein	

---

**Other function**


---

**Dehydrogenase activity**

DPYD	dehydropyrimidine dehydrgenase	GLT1
------	--------------------------------	------

**Ligase activity**

MARS	methionine-tRNA synthetase	MES1
------	----------------------------	------

**Peptidase activity**

SEC11L1	signal peptidase complex 18Kda	SEC11
---------	--------------------------------	-------

**Receptor activity**

REEP6	receptor accessory protein 6	
-------	------------------------------	--

**Translation elongation factor activity**

EEF1B2	eukaryotic translation elongation factor 1 beta 2	
--------	---	--

**Translation initiation factor activity**

ITGB4BP	integrin beta 4 binding protein	TIF6
---------	---------------------------------	------

---

**Unknown function**


---

EBNA1BP2	EBNA1 binding protein 2	EBP2
GPIAP1	GPI-anchored protein p137 precursor	
HDCMA18P	hypothetical protein DKFZp564K112.1 (HDCMA18P)	
LOC389217	similar to SET protein (Phosphatase 2A inhibitor I2PP2A) (I-2PP2A) (Template activating factor I) (TAF-I) (HLA-DR associated protein II) (PHAPII) (Inhibitor of granzyme A-activated DNase) (IGAAD)	
MGC3731	hypothetical protein LOC79159	
NOC2L	nucleolar complex associated 2 homolog (S. cerevisiae; hypothetical protein DKFZp564C186.1)	NOC2
NOC3L	nucleolar complex associated 3 homolog (S. cerevisiae)	NOC3
NOL10	nucleolar protein 10 (hypothetical protein FLJ14075)	ENP2

NOLA3	nucleolar protein family A, member 3	NOP10
PES1	Pescadillo homolog 1 containing BRCT domain	NOP7
RBM12B	RNA binding motif protein 12B	
RP13-36C9.1	cancer/testis antigen CT45-2	
SYNGR2	synaptogyrin 2	

---

---

**Tabella 10. Identificazione dei ligandi presenti nel canale di controllo GST**


---

<b>Gene</b>	<b>Proteina</b>
ACTB	Beta actin
AK3	Adenylate kinase
ANK3	Ankyrin-3
ANXA11	Annexin A11
BXDC2	BRIX
DCD	Dermicidin precursor sor
DDB1	Damage-specific DNA binding protein 1
DMXL2	Dmx-like 2
DNAH3	dynein, axonemal, heavy polypeptide 3
EEF1G	pancreatic tumor-related protein
EGFR	ADP-ribosylation factor 4
F8	Coagulation factor VIII
GIMAP4	IgM heavy chain V region (fragment)
HBZ	Zeta globin
HEJ1	HEJ 1
HSPA6	Heat shock 70 KDa protein 6
HSPCB	Heat shock 90 KDa protein
IL1F8	Interleukin I family, isoform 1
ITGAX	Integrin alpha X precursor
K-ALPHA-1	Alpha tubulin
KRT1	Keratin1
KRT14	Keratin 14
KRT19	unnamed protein product
KRT9	cytokeratin 9
KTR10	keratin 10
KTR15	unnamed protein product
LRP4	Low density lipoprotein receptor-related protein 4
MYH6	Myosin heavy chain, cardiac muscle alpha isoform
RGD1307877	Glial fibrillary acidic protein
RP3-402G11	Selenoprotein
SPATA	spermatogenesis associated factor protein
TSR1	hypothetical protein LOC55720
TUBB4	Beta-tubulin
WBP4	Formin binding protein 3
ZNF561	Zinc finger protein 56

---

Tabella 11. Analisi per spettrometria di massa del canale GST-Rps19

Gene	Sequenza peptidica identificata	Errore (ppm)	Mascot Score	Copertura di Sequenza
<b>NTPase activity</b>				
GTPBP4	LPTIDPNTR	107	38	12%
	LALGQINIAK	125	38	
	QSLEYLEQVR	138	38	
	TLLLCGYPNVGK	147	50	
	ADVVDVQPYAFTTK	149	40	
PSMC5	GVCTEAGMYALR	89	38	20%
	VDPLVSLMMVEK	123	71	
	GVCTEAGMYALR	150	65	
	VPDSTYEMIGGLDK	110	50	
	VHVTQEDFEMAVAK	100	40	
	TMLELLNQLDGFEATK	125	38	
PSMC6	EMFNYAR	91	38	1%
RAB11B	DHADSNIVIMLVGNK	112	42	7%
XAB1	SMSLVLDEFYSSLR	120	83	4%
<b>Hydrolase/helicase activity</b>				
DDX5	QVSDLISVLR	107	40	10%
	APILIATDVASR	88	45	
	TTYLVLDEADR	96	50	
	RLMEEIMSEKENK	140	39	
	TGTAYTFFTPNNIK	110	46	
DDX17	APILIATDVASR	98	38	2%
DDX18	NGTGVLLSPTR	12	53	9%
	EPLYVGVDDDK	150	44	
	LGNGINIIVATPGR	50	78	
	SAEAQKLGNGINIIVATPGR	144	38	
DDX21	DFSDITKK	112	38	42%
	LHGELQDR	140	38	
	TIIFCETK	150	38	
	IGVPSATEIHK	150	38	
	EQLGEEIDSK	150	39	
	APQVLVLAPTR	133	40	
	TAITVEHLAIK	106	38	
	NGIDILVGTPGR	147	48	
	GVTFLFPIQAK	150	38	
	STYEQVDLIGK	122	38	
	TFSFAIPLIEK	110	38	
	RIGVPSATEIHK	128	38	
	EAQELSONSAIK	136	45	
	EQLGEEIDSKVK	139	39	
	GRAPQVLVLAPTR	140	48	
	STYEQVDLIGKK	142	40	

	EEYQLVQVEQK	150	38	
	NGSFGVLVATNVAAR	147	40	
	LLDSVPPTAISHFK	109	38	
	EGAFSNFPISEETIK	138	49	
	NEEPSEEEIDAPKPK	127	40	
	WQLSVATEQPELEGPR	111	40	
	TFHHVYSGKDLIAQAR	105	38	
	LLDSVPPTAISHFKQSAEK	148	38	
	GLDIPEVDLVIQSSPPKDVESYIHR	109	40	
DDX24	DKLDILGAAETGSGK	95	76	2%
DDX3X	VGSTSENIQK	146	38	4%
	VRPCVVYGGADIGQQIR	147		40
DDX41	SGNTGIATTFINK	150	44	2%
DDX50	VLVLAPTR	111	44	5%
	VLVATNVAAR	145	38	
	LSSNAVSQITR	107	50	
	TFSFAIPLIER	57	48	
DDX54	LLVEFAR	147	41	27%
	TSFFLVR	141	48	
	ELALQTLK	144	39	
	TIPVILDGK	129	48	
	AGLTEPVLIR	146	39	
	ATIFEINASSR	124	74	
	LVHVAVEMSLK	117	47	
	SGGFQSMGLSYPVFK	74	40	
	LPGGHQTVLFSATLPK	144	38	
	LQSVEYVVFDEADR	78	45	
	TIPVILDGKDVVAMAR	106	38	
	EMDLVGLGLHPLFSSR	150	38	
	LFEMGFAEQLQEIIAR	84	69	
	AKEMDLVGLGLHPLFSSR	96	38	
	GLDIPLLDNVINYSFPAK	112	38	
	MEDQFAALHENPDIIIATPGR	5	54	
	VPQSVVDEEDSGLQSTLEASLELR	85	49	
DHX9	LGGIGQFLAK	119	40	21%
	MLNMIR	105	43	
	LNMATLR	113	38	
	MGEEAEIR	115	38	
	MTPSYEIR	106	38	
	LSMSQLNEK	96	40	
	DFVNYLVR	146	38	
	VFDPPVPGVTK	138	38	
	DVVQAYPEVR	59	40	
	TPLHEIALSIK	122	38	
	AAMEALVVEVTK	133	38	
	KVFDPVPGVTK	93	39	
	LAAQSCALSLVR	146	40	
	YPSPFFVFGEK	127	38	
	SFIAEMTIYIK	120	38	

	ETPFELIEALLK	135	39	
	YQILPLHSQIPR	143	38	
	GISHVIVDEIHER	106	38	
	QPAIISQLDPVNER	150	40	
	VQSDGQIVLVDDWIK	111	38	
	ELDALDANDELTPGR	52	38	
	GMTLVTPQLQLLFASK	109	38	
	QLYHLGVVEAYSGLTK	96	42	
	AIEPPPLDAVIEAEHTLR	30	38	
	KFESEILEAISQNSVVIIR	3	40	
	NELMYQLEQDHDLDLQAILQER	7	38	
	LGGIGQFLAKAIEPPPLDAVIEAEHTLR	150	38	
	DINTDFLLVVLR	142	38	
DHX15	EVDDLGPVEVDIK	148	54	2%
DHX36	ELDILLQEK	146	38	3%
	NLQSDVLMTVVK	108	63	
	ASLLDDYQLPEILR	110	69	
MCM2	VAVGELTDEDVK	149	38	6%
	QLVAEQVTYQR	147	49	
	DTVDPVQDEMLAR	112	53	
	YDPSLTFSENVDLTEPIISR	148	81	
MCM6	YLQLAEELIRPER	147	46	7%
	EIESEIDSEELINK	17	76	
	LFLDFLEEFQSSDGEIK	143	45	
MCM7	AGILTTLNAR	150	38	4%
	SITVLVEGENTR	147	47	
RUVBL2	TTEMETIYDLGTK	20	70	3%
SKIV2L2	ALFATETFAMGINMPAR	110	38	2%
SMARCA5	FEYLLK	147	48	14%
	EILFYR	139	43	
	FDWFLK	116	39	
	DIDILNSAGK	147	44	
	YLVIDEAHR	117	52	
	LDSIVIQQGR	106	61	
	FITDNTVEER	33	64	
	KANYAVDAYFR	83	62	
	TLQTISLLGYMK	130	45	
	LRLDSIVIQQGR	122	38	
	LGFDKENVYDELRL	150	38	
	ESEITDEDIDGILER	121	70	
	TPEEVIEYSVFWER	43	62	
	NFTMDTESSVYNFEGEDYR	34	61	
XRN2	TGGYLTESGYVNLQR	88	38	6%
	ELTMASLPFTFDVER	115	38	
	AALEEVYVYDLTPEETR	102	38	
	VQMIMLAVGEVEDSIFK	105	35	

---

**Isomerase activity**


---

DKC1	LHNAIEGGTQLSR	140	41	6%
------	---------------	-----	----	----

	ALELTGALFQRPPPIAAVK	151	38	
PPIH	IIDGLLVMR	105	61	10%
	IELFADVVPK	115	69	
<b>Kinase activity</b>				
CSNK2A1	ALDYCHSMGIMHR	112	55	4%
SRP72	LTNAEGVEFK	144	43	10%
	VLANNLSFEK	147	38	
	GTQGATAGASSELDASK	135	63	
	TVSSPPTSPRPGSAATVSASTSNIIPPR	146	38	
PRKCQ	AKGQSLQDPFLNALR	6	41	2%
<b>Splicing factor activity</b>				
SF3B1	TEILPPFFK	76	43	8%
	QLVDTTVELANK	145	42	
	LLVDVDESTLSPEEQK	122	38	
	ILVVIEPLIIDEDYYAR	113	40	
	MVMETIEKIMGNLGAADIDHK	107	38	
	AAGLATMISTMRPDIDNMDEYVR	109	55	
SF3B2	YGPPPSYPNLK	150	40	3%
	GIEKPPFELPDFIK	150	38	
SF3B3	SVAGGFVYTYK	126	55	11%
	LTISSPLEAHK	147	39	
	IVPGQFLAVDPK	93	38	
	FLAVGLVDNTVR	123	53	
	ILELLRPDPNTGK	150	38	
	TVLDPVTGDLSDTR	107	67	
	ITLETDEDMVTEIR	42	90	
	NENQLIIFADDTYPR	83	46	
	ELAAEMAAAFNLNENLPESIFGAPK	114	60	
SFRS9	HGLVPFAFVR	144	44	10%
	IYVGNLPTDVR	24	38	
SFRS10	VDFSITK	150	31	12%
	YGPIADVSIYDQQSR	133	85	
<b>Structural constituent of Ribosome</b>				
RPL10A	DTLYEAVR	20	46	10%
	KYDAFLASESLIK	74	50	
RPL14	GTAAAAAAAAAAAAAK	60	96	18%
	VAYVSFGPHAGK	113	50	
	LVAIVDVIDQNR	15	68	
RPL24	VFQFLNAK	148	52	13%
	TAMAAAKAPTK	97	70	
RPL27A	TGAAPIIDVVR	99	71	7%
RPL3	HGSLGFLPR	20	38	2%
RPL4	NIPGITLLNVSK	142	44	3%
RPL6	YYPTEDVPR	147	51	3%
RPL7	ASINMLR	145	38	31%
	KVLQLLR	138	38	

	SVNELIYK	145	38	
	EVPAVPETLK	129	38	
	SVNELIYKR	71	38	
	IALTDNALIAR	76	47	
	AGNFYVPAEPK	92	38	
	KAGNFYVPAEPK	5	42	
RPL7A	AGVNTVTTLVENK	148	76	7%
RPL8	AVVGVVAGGGR	85	61	5%
RPL9	TILSNQTVDIPENVDTLK	90	49	10%
RPLP0	EDLTEIR	131	38	2%
RPLP1	AAGVNVPEPFWPGLFAK	117	41	14%
RPLP2	LASVPAGGAVAVSAA	50	38	31%
RPS10	IAIYELLFK	135	64	5%
RPS14	IEDVTPIPSDSTR	14	46	9%
RPS16	DILIQYDR	122	38	40%
	ALVAYYQK	90	38	
	TATAVAHCK	84	38	
	GPLQSVQVFGR	75	38	
	VNGRPLEMIEPR	150	38	
	GGGHVAQIYAIR	95	38	
RPS2	GTGIVSAPVPK	68	60	11%
RPS23	VANVSLALYK	67	87	8%
RPS24	QMVIDVLHPGK	54	69	22%
	TTGFGMIYDSLIDYAK	60	70	
RPS26	GHVQPIR	100	70	17%
	LHYCVSCAIHSK	98	49	
RPS3	ELTAVVQK	34	38	30%
	KFVADGIFK	68	42	
	TEIILATR	150	57	
	AELNEFLTR	54	49	
	ELAEDGYSGVEVR	56	57	
	KLPDPHVSIVEPK	144	43	
	DEILPTTPISEQK	41	38	
RPS4X	LSNIFVIGK	131	55	18%
	YALTGDEVK	90	34	
	TIRYPDPLIK	150	33	
	TDITYPAGFMDVISIDK	5	55	
RPS5	QAVDVSPLR	10	44	4%
RPS6	LIEVDDER	20	56	8%
	DIPGLTDTTVPR	39	61	
RPS7	AIIIFVPVQLK	60	41	6%
RPS8	LTPEEEEEILNK	8	47	5%
RPSA	LLVVTDPK	150	50	25%
	SDGIYIINLK	70	52	
	FAAATGATPIAGR	33	51	
	AIVAIENPADVSVISS	46	70	
	FLAAGTHLGGTNLDFQMEQYIYK	15	38	
RSL1D1	FFTTPSK	72	38	33%
	SPNPSTPR	128	40	

LLFVKTEK	41	38
QLDKEQVR	10	38
RLLPSLIGR	19	38
VPVSVNLLSK	5	42
AVDALLTHCK	18	38
KVPVSVNLLSK	84	44
TVSQIISLQTLK	37	38
DDVAPESGDTTVK	8	43
TVSQIISLQTLKK	41	38
LLSSFDFFLTDAR	26	100
EINDCIGGTVLNISK	110	50
ATNESEDEIPQLVPIGK	54	38
DEPNSTPEKTEQFYR	5	39

Transcription factor				
BAZ1B	GGLGYVEETSEFEAR	150	75	1%
HNRPD	IFVGGLSPDTPEEK	74	78	12%
	EYFGGFGEVESIELPMDNK	9	74	
ILF2	VLQSALAAIR	67	90	25%
	KLDPELHLDIK	150	38	
	ILPTLEAVAALGNK	70	77	
	VKPAPDETSFSEALLK	23	69	
ILF3	INNVIDNLIVAPGTFEVQIEEVR	150	38	
	AQDPSEVLTMLTNETGFEISSDATVK	150	51	
	AYAALAALEK	118	63	14%
	LFPDTPLALDANK	127	83	
	VADNLAIQLAAVTEDK	36	63	
	EPPLSLTIHLTSPVVR	78	52	
NKRF	VLAGETLSVNDPPDVLDR	89	55	
	VADNLAIQLAAVTEDKYEILQSVDDAAIVIK	57	46	
	DIEQIIR	131	44	13%
	EIPPADIPK	150	42	
	INITYMLTR	92	43	
PURA	LLTDGYACEVR	131	52	
	TNPEYIYAPLK	84	54	
	VILESEVIAEAVGVK	130	64	
TAF15	FFFVVGSNK	94	65	11%
	LIDDYGVEEPAELPEGTSLTVDNK	33	59	
TRIM28	GEATVSFDDPPSAK	150	44	2%
UBTF	LSPPYSSPQEFAQDVGR	139	46	2%
	DYEVELLR	122	39	13%
	KKDYEVELLR	148	38	
	HPELNISEEGITK	82	40	
XPO5	ITLTELILDAQEHVK	143	38	
	DPLLLAIIPK	145	45	3%
	LFSSVTFETVEESK	62	61	
YBX1	QGETQTELVMFILLR	62	70	
	GAEAAANVTGPGGVPVQGSK	53	57	10%
	SVGDGETVEFDVVEGEK	5	105	

<b>Transferase activity</b>				
FDFT1	ALDTLEDDMTISVEK	86	71	6%
	AIIYQYMEEIYHR	140	94	
FTSJ3	FQFLQK	150	38	13%
	EVEVQAK	102	54	
	TSVTDFLR	150	48	
	AANPVDFLSK	141	42	
	DLIDNSFNR	128	38	
	AEAVVNTVDISER	150	53	
	ILDPEGLALGAVIASSK	78	79	
	SDDDGFEIVPIEDPAK	51	38	
	LTEVQDDKEEEEEENPLLPLEEK	138	38	
	NAT10	SMDLSEYIIR	150	
	LDYLGVSYGLTPR		41	
	LGQAELVVIDEAAAIPPLVK		41	
NOL1	GVNLDPLGK	119	39	14%
	DLAQALINR	137	60	
	IQDIVGILR	145	51	
	GADSELSTVPSVTK	137	38	
	VLLDAPCSGTGVISK	46	44	
	LGVTNTIISHYDGR	131	50	
	ELLSAIDSVNATSK	30	78	
	LVPTGLDFGQEGFTR	72	53	
	SPEAKPLPGKLPKGAVQTAGK	137	47	
	ZC3HAV1	ASLEDAPVDDLTR	150	
<b>Transporter activity</b>				
COPA	MHSLLIK	63	38	13%
	VWDISGLR	105	38	
	TALNLFFK	147	44	
	GFFEGTIASK	150	47	
	VLTIDPTEFK	106	51	
	TLDLPIYVTR	126	56	
	NLSPGAVESDVR	105	55	
	EYIVGLSVETER	86	55	
	DADSITLFDVQKQK	54	65	
	SILLSVPLLVDNK	126	61	
	GITGVDLFGTTDAVVK	113	51	
	VTTVTEIGKDVIGLR	69	59	
	LLELGPKPEVAQQTR	71	56	
	ASNLENSTYDLYTIK	8	43	
	COPB2	GSNVALGYDEGSIIVK	121	
CSE1L	TGNIPALVR	128	50	14%
	VIVPNMEFR	149	62	
	SANVNEFPVLK	144	58	
	DAAIYLVTSLASK	126	49	
	ALTLPGSSENEYIMK	60	52	
	LVLDAFALPLTNLFK	66	38	

	YGALALQEIFDGIQPK	68	53	
	AADEEAFEDNSEEYIR	42	66	
	LLQTDDEEEAGLLELLK	98	86	
IPO4	QGCTVAEK	121	67	2%
	LLMASPTR	90	65	
IPO7	ETENDDLTNVIQK	110	64	3%
	ENIVEAIIHSPYLIR	117	68	
NPEPL1	YHAAVLNTSAEWAAACVK	58	44	3%
STAU1	VSVGFEVGEGERK	38	43	3%
SSR4	FFDEESYSLLR	112	69	6%
XPO1	YVVGIIK	121	38	7%
	IYLDMLNVYK	148	45	
	EFAGEDTSDLFLEER	40	84	
	LLSEEVDFSSGQITQVK	46	62	
	MAKPEEVLVENDQGEVVR	14	40	

---

**DNA/RNA/protein binding capacity**


---

AATF	DKGGPEFSSALK	150	41	2%
ACTR1B	DWNDMER	120	67	2%
C1orf77	LGRPIGALAR	91	69	10%
CCT2	GATQQILDEAER	112	58	2%
CCT8	DIDEVSSLLR	118	46	2%
CEBPZ	ALLVQVVNK	133	40	10%
	EQIDTLFK	127	39	
	QTLRLRPGGK	122	38	
	MLSALLTGVR	148	38	
	QAMFLNLVYK	65	39	
	LYQHEINLFK	150	38	
	ELLITDLLPDNR	136	40	
	DKQNIFEFFER	109	40	
	EESQIPVDEVFFHR	69	38	
	KLETEETVPETDVETK	59	40	
CENPC1	VSDEEDK	58	80	3%
	QMPVVGSK	89	55	
	ILATDVSSK	95	95	
COPG	SIATLAITLLK	148	69	5%
	SLEELPVDIILASVG	32	51	
	LLLLDTVTMQVTAR	126	95	
DHX30	IPQLLER	100	51	17%
	ALTQFPLPK	120	39	
	VPGFMYPVK	123	38	
	EYLTTLGQR	92	50	
	AVAGWEEVLR	148	76	
	LQSDDILPLGK	146	48	
	AIFQQPPVGVV	99	38	
	TPLENLVLQAK	98	47	
	EHYLEDILAK	130	38	
	ATISLSDSLLR	77	53	
	MVPFQVPEILR	89	40	

	DVNTDFLLILLK	128	72	
	GVLMAGLYPNLIQVR	90	44	
	AVDEAVILLQEIGVLDQR	121	55	
	DSGPLSDPITGKPYVPLLEAEEVR	23	38	
	WQDRSSRENYLEENLLYAPSLR	132	38	
DNAJC9	ELGLDEGVDSLK	150	41	8%
	ISLEDIQAFEK	150	43	
FBL	TNIIPVIEDAR	120	62	11%
	NLVPGESVYGEK	150	38	
	VSISEGDDKIEYR	150	52	
FUSIP1	DAEDALHNLDR	8	57	6%
GNB2L1	DVLSVAFSSDNR	99	66	4%
HIST1H1C	ALAAAGYDVEK	63	42	12%
	SETAPAAPAAPAPAEK	40	40	
HIST1H1D	ALAAAGYDVEK	72	46	10%
	ASGPPVSELITK	125	57	
HIST1H2AK	NDEELNK	110	38	5%
HIST1H2BL	LLLPGELAK	112	38	7%
HIST1H2BO	QVHPDTGISSK	123	50	17%
	AMGIMNSFVNDIFER	90	53	
HIST2H4	ISGLIYEETR	93	65	36%
	DNIQGITKPAIR	136	48	
	TVTAMDVVYALK	141	80	
	TVTAMDVVYALKR	142	38	
HNRPA2B1	DYFEEY GK	28	38	8%
	TLETVPLER	75	40	
	IDTIEIITDR	76	47	
HNRPA3	LFIGGLSFETTDDSLR	45	94	5%
HNRPC	VPPPPPIAR	22	58	9%
	SDVEAIFSK	70	51	
	KSDVEAIFSK	62	61	
HNRPDL	DLTEYLSR	70	46	3%
HNRPF	SSQEEVR	70	85	13%
	TEMDWVLK	98	90	
	VHIEIGPDGR	110	56	
	DLSYCLSGMYDHR	123	48	
	QSGEAFVELGSEDDVK	57	54	
HNRPR	LFVGSIPK	150	43	8%
	ENILEEFSK	150	46	
	TGYTLDVTTGQR	30	57	
	NLATTVTEEILEK	116	79	
	DLYEDELVPLFEK	12	48	
HNRPU	FIEIAAR	89	38	5%
	GYFEYIEENK	100	41	
	YNILGTNTIMDK	141	38	
	NFILDQTNVSAAAQR	78	97	
HNRPUL2	EEAYHSR	112	46	2%
	ANFSLPEK	109	38	
HP1BP3	SGASVVAIRK	101	40	9%

	GASGSFVVVQK	65	40	
	YIIHKYPSLELER	147	47	
	SSAVDPEPQVK	112	38	
IGF2BP1	MVIITGPPEAQFK	150	39	10%
	LLVPTQYVGAIGK	149	50	
	TVNELQNLTAAEVVVPR	10	54	
	LYIGNLNESVTPADLEK	21	50	
IGF2BP3	ALQSGPPQSR	10	38	
	IPVSGPFLVK	5	38	
	QKPCDLPLR	9	51	
	DQTPDENDQVVVK	9	43	
	SITILSTPEGTSAACK	6	66	
IMP3	LYALGLVPTR	140	52	5%
ITGB4BP	ETEEILADVLK	31	63	4%
LYAR	FQNWMK	150	78	3%
	QQAWIQK	95	48	
NCL	ALELTGLK	145	38	17%
	TGISDVFAK	124	42	
	GIAYIEFK	131	41	
	NDLAVVDVR	110	68	
	EVFEDAAEIR	150	73	
	GFGFVDFNSEEDAK	147	38	
	FGYVDFESAEDLEK	73	75	
	GLSEDTTEETLKESFDGGSVR	91	43	
	TLVLSNLSYSATEETLQEVFEK	28	67	
NIP7	LHVTALDYLAQYK	142	62	10%
NOLA1	FYIDPYK	105	38	13%
	VDEIFGQLR	18	38	
	VPYFNAPVYLENK	6	49	
NOL5A	VVSLSEYR	148	40	1%
PABPC3	IVATKPLYVALAQR	150	56	2%
PAK1IP1	GEQYVVIIQNK	142	61	12%
	FLSESVLAVAGDEEVIR	103	77	
	IDIQDLTASISGTITNEK	14	83	
PNN	LLALSGPGGGR	140	38	5%
	IEFAEQINK	120	42	
	LTEVPVEPVLTVHPESK	117	42	
PPP2R1A	LSTIALALGVER	120	50	2%
RAB1B	QWLQEIDR	98	49	4%
RAP1B	LVVLGSGGVGK	85	40	6%
RBM19	NLPYTSTEEDLEK	117	38	5%
	ILGENEEEEEDLAESGR	146	92	
	VLLPEGGITAIVEFLEPLEAR	127	64	
RBMX	ALEAVFGK	72	50	13%
	IVEVLLMK	137	59	
	LFIGGLNTETNEK	91	57	
	GFAFVTFESPADAK	60	61	
RNPC2	IESIQLMMDSETGR	116	36	8%
	TDASSASSFLDSDELER	16	94	

SART3	IQLIFER	150	38	4%
	SALQALEMDR	148	44	
	EFESAIVEAAR	133	85	
	LAEQAYIDFEMK	12	60	
SNRPA1	SLTYLSILR	94	46	4%
SNRPG	HVQGILR	80	42	9%
SNRPN	VLGLVLLR	104	48	10%
	GENLVSMTEGPPPK	19	52	
SRP68	ALLQQPEDDSKR	133	39	2%
SURF6	LLQEALK	65	48	2%
SYNCRIP	LFVGSIPK	120	44	2%
	TGYTLDVTTGQR	30	57	

---

**Other function**


---

*Dehydrogenase activity*

DPYD	VKEALSPIK	120	43	1%
------	-----------	-----	----	----

*Ligase activity*

MARS	ITQDIFQQLK	148	37	5%
	GFVLQDTVEQLR	145	62	
	FFGGYVPEMVLTPDDQR	5	40	

*Peptidase activity*

SEC11L1	VGEIVVFR	107	70	4%
---------	----------	-----	----	----

*Receptor activity*

REEP6	NVKPSQTPQPK	150	77	6%
-------	-------------	-----	----	----

*Translation elongation factor activity*

EEF1B2	SPAGLQVLNDYLADK	12	82	15%
	YGPADVEDTTGSGATDSK	28	74	

---

**Unknown function**


---

EBNA1BP2	DLEWVER	95	38	14%
	QAQAAVLAVLPR	150	40	
	LDVTLGPVPEIGGSEAPAPQNK	120	39	
GPIAP1	QILGVIDKK	150	38	1%
HDCMA18P	MGEEVIPLR	149	39	4%
	SSAVVELDLEGTR	147	38	
LOC389217	LSQMQNK	110	62	1%
MGC3731	DPLLSQR	120	72	3%
NOC2L	QLAIHLR	60	45	4%
	EIQLEISGK	106	38	
	LEDLNFPEIK	105	46	
NOC3L	LGQASLGVIK	142	76	8%
	SPLLPVAVLEGLAK	137	54	
	FYLENLEQMVK	77	38	
	YSSEVATESPLDFTK	84	40	
	SMLMEQDPDVAVTVR	6	47	
NOL10	QLTFTLKR	150	40	2%
	LLEQQELR	146	38	
NOLA3	VLMTQQPRPVL	139	45	17%
PES1	GSATNYITR	20	56	1%

---

RBM12B	YAFVMFK	145	38	16%
	GVGLGEALVK	149	44	
	NFPFDVTK	144	38	
	NLSLSIDER	124	44	
	FLGTEVLLR	150	67	
	AENPYLFLR	142	41	
	DSSVELFLSSK	148	77	
	GLPYLVNEDDVR	145	56	
	LLGLPFIAGPVDIR	108	42	
	DPPIYSVGAFENFR	67	42	
	FFADFLAEDDIYLLYDDK	90	40	
	FLYKDNRTR	43	40	
RP13-36C9.1	VAVDPETVFK	144	53	13%
	IFEMLEGVQGPTAVR	12	58	
SYNGR2	AGGSFDLR	90	42	4%

---

Tabella 12. Analisi per spettrometria di massa del canale di controllo GST

Gene	Sequenza peptidica identificata	Errore (ppm)	Mascot Score	Copertura di Sequenza
ACTB	DLTDYLMK	4	38	14%
	QEYDESGPSIVHR	53	40	
	EITALAPSTMK	85	38	
	DLYANTVLSGGTTMYPGIADR	5	38	
BXDC2	ILIFSSR	132	48	15%
	FVLNLIK	121	50	
	LFVINEVCEMK	90	57	
	FLVQNIHTLAELK	105	38	
	NFQIIEEDAALVEIGPR	142	38	
DCD	ENAGEDPGLAR	35	40	20%
	DAVEDLESVGK	43	38	
DDB1	LFMLLLEK	114	50	1%
EEF1G	KLDPGSEETQTLVR	52	38	11%
	LDPGSEETQTLVR	21	38	
	ALIAAQYSGAQVR	14	38	
	QAFPNTNR	53	38	
EGFR	MLLVDELK	120	38	1%
	ILMVGLDAAGK	90	53	
HBZ	ISTQADTIGTETLER	67	88	10%
HEJ1	EQYSAVIAK	26	41	15%
HSPA6	TTPSYVAFTDTER	150	41	2%
HSPCB	DNSTMGYMMAK	116	75	1%
ITGAX	ESHVAMHR	110	45	2%
	IAPPASDFLAHIQK	80	40	
K-ALPHA-1	DVNAAIATIK	110	39	2%
KRT1	SLDLDSIIAEVK	150	52	4%
	WELLQQVDTSTR	150	51	
KRT14	QFTSSSSMK	85	94	8%
	VTMQNLNDR	63	73	
	MSVEADINGLR	142	80	
	CEMEQQNQEYK	120	47	
KRT19	VLDELTAR	104	54	2%
KRT9	MTLDDFR	150	38	8%
	TLLDIDNTR	104	40	
	QEYEQLIAK	141	43	
	QVLDNLTMEK	69	38	
	VQALEEANNDLENK	2	54	
KTR10	DAAEFNEK	90	91	5%
	LENEIQTYR	123	46	
	ALEESNYELEGK	107	83	
KTR15	IMATTIDNSR	72	107	5%
	QGVEADINGLR	80	50	
RGD1307877	LALDIEIATYR	1	74	3%

---

SPATA	AVAPSIIFDELDAVER	72	89	7%
TSR1	QIDAPGDPFLNPR	68	70	7%
	LLHIVGYGDFQMK	72	38	
	LEEMFPDEVDTPR	63	42	
	LLLLDTQQEAGMLLR	46	83	
TUBB4	LAVNMVPFPR	74	40	10%
	ISVYYNEATGGK	53	39	
	IMNTFSVVPSPK	89	38	
	AILVDLEPGTMDSVR	123	41	
ZNF561	SFTNFSQLYAPVK	110	130	3%

---

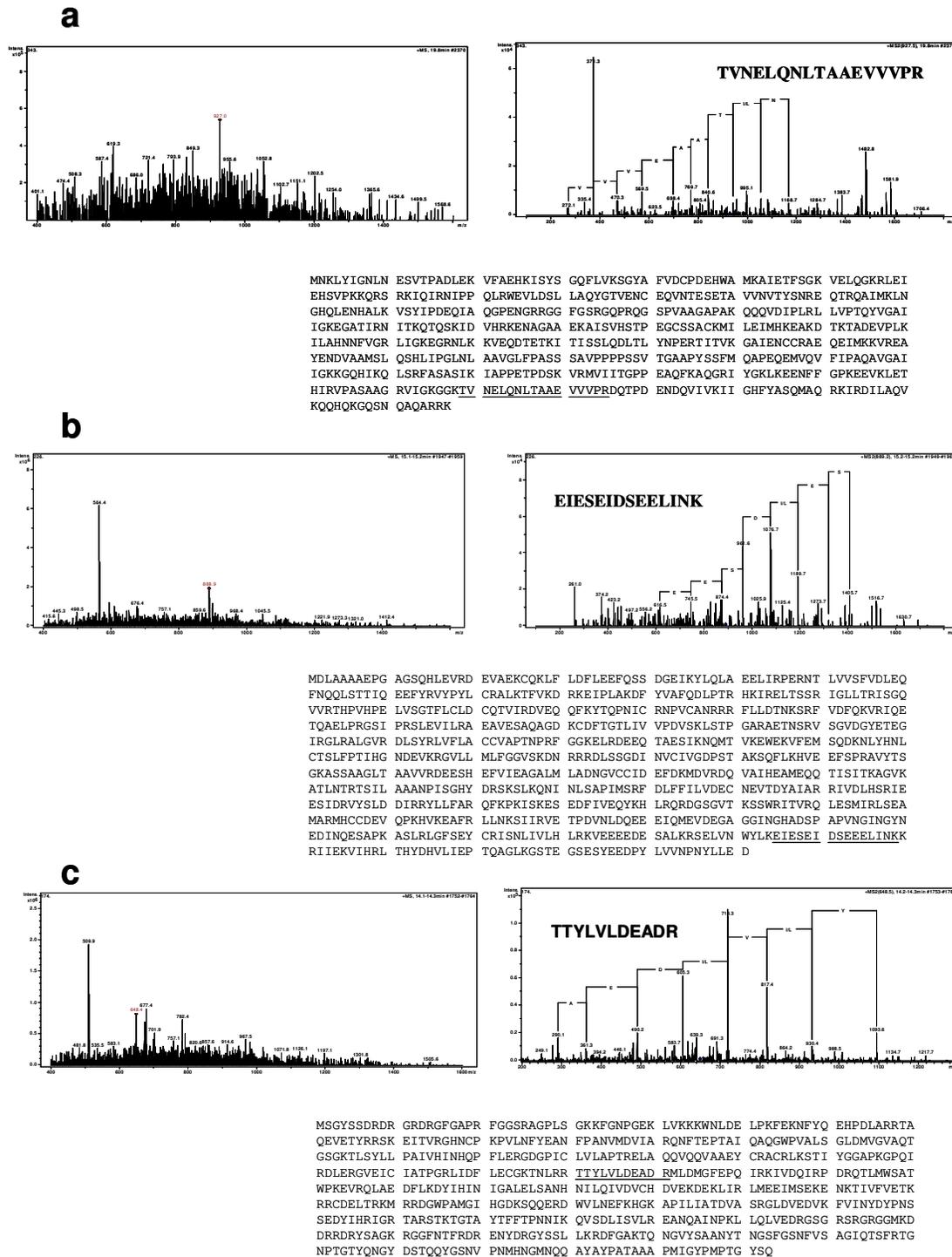
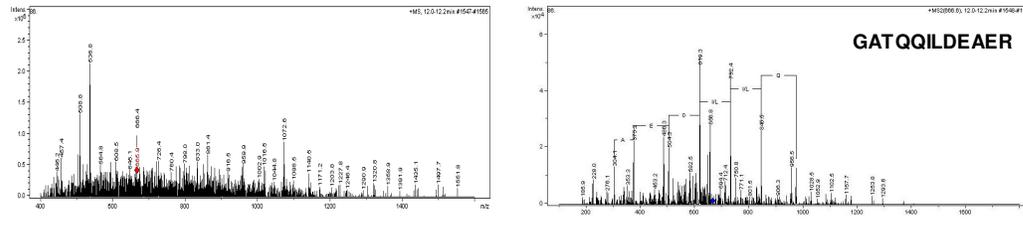
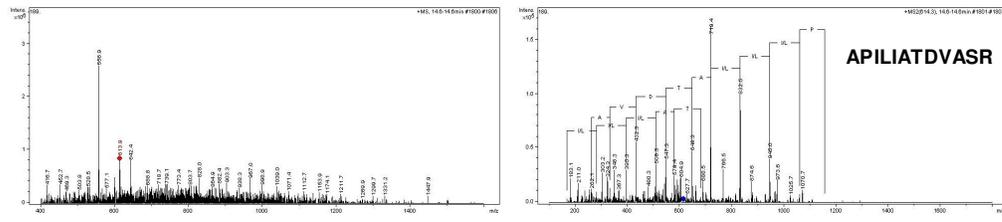


Figura 21. Spettro di massa, spettro di frammentazione e sequenza aminoacidica identificata per le proteine IGF2BP1 (a), MCM6 (b), DDX5 (c).



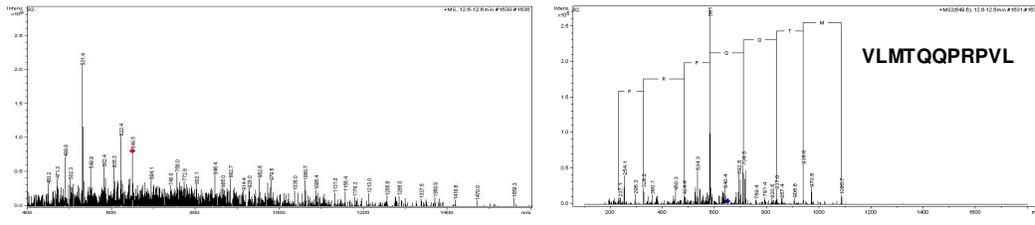
**a**

ASLSLAPVNI FKAGADEERA ETARLTSFIG AIAIGDLVKS TLGPKGMDKI LLSSGRDASL MVTNDGATIL  
 KNIGVDNPAA KVLVDMRSRQ DDEVGDGTTT VTLAAELLR EAESLI AKKI HPQTIIAGWR EATKAAREAL  
 LSSAVDHGSD EVKFRQDLMN IAGTTLSSKL LTHHKDHFTK LAVEAVLRK GSGNLEAHI IKKLGSSLAD  
 SYLDEGLLD KKIGVNQPKR IENAKILIAN TGMDDKIKI FGSRVRVDST AKVAIEHA EKEKMEKVER  
 ILKHGICFI NRQLIYNYPE QLFGAAGVMA IEHADFAGVE RLALVTGGEI ASTFDHPELV KLGSCKLEE  
 VMIGEDKLIH FSGVALGEAC TIIVLRGATQQ ILDEAERSLH DALCVLAQTV KDSRTVYGGG CSEMLMAHAV  
 TQLANRTPGK EAVAMESYAK ALRMLPTIIA DNAGYDSADL VAQLRAHSE GNTTAGLDMR EGTIGDMAIL  
 GITESFQVKR QVLLSAAEAA EVILRVDNII KAAPRKRVPD HHPG



**b**

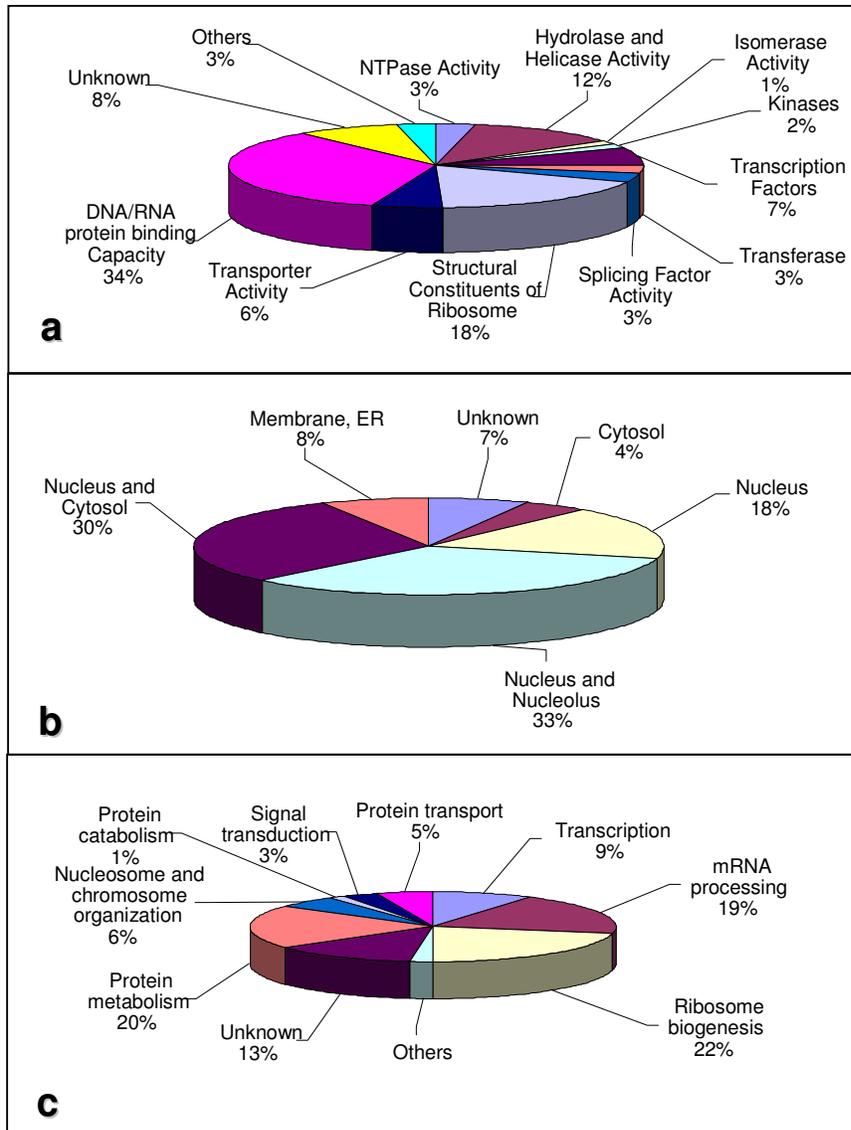
MRGGGFGRDR RDRDRGGFGA RGGGLPPKK FGNPGERLRK KKWDLSELPK FEKNFYVEHP EVARLTPYEV  
 DELRRKKEIT VRGGDVCPKP VFAPHHANFP QYVMDVLMQ HFTEPTPIQC QGFPLALSGR DMVGIAQTGS  
 GKTLAYLLPA IVHINHQPYL ERGDGPICLV LAPTRELAQQ VQVADDYQK CSRKSTCIY GGAPKGPQIR  
 DLERGVEICI ATPGRLIDFL ESGKTNLRRC TYLVLDEADR MLDMGFEPQI RKIVDQIRPD RQTLMWSATW  
 PKEVRLAED FLRDYQINLV GNLELSANHN ILQIVDVCME SEKDHKLIQL MEEIMAEKEN KTIIFVETKR  
 RCDLIRMR RDGNPAMCIH GDKSQPERDW VLNEFRSGKA PILIATDVAS RGLDVEDVKV VINYDYPNS  
 EDYVHRICRT ARSNKGTAY TFFTPCNLQ ARELIKVLEE ANQAINPKLM QLVDRGCGG GCGRSRYRT  
 TSSANNPLM YQDECDRLR GVKDGGRRDS ASYRDRSETD RAGYANGSGY GSPNSAFGAQ AGQYTYGQGT  
 YGAAAYTSS YTAQEYAGT YGASSTTSTG RSSQSSSQF SGIGRSQQQP QPLMSQQFAQ PPGATNMIGY  
 MGQYAYQYP PPPPPPPSRK



**c**

MFLQYLNEQ GDRVYTLKFK DPMQQTCSA HPARFSPDDK YSRHRTIKK RPKVLTQQP RPVL

**Figura 22.** Spettro di massa, spettro di frammentazione e sequenza amminoacidica per le proteine CCT2 (a), DDX1 (b), NOLA3 (c), identificate con un singolo peptide.



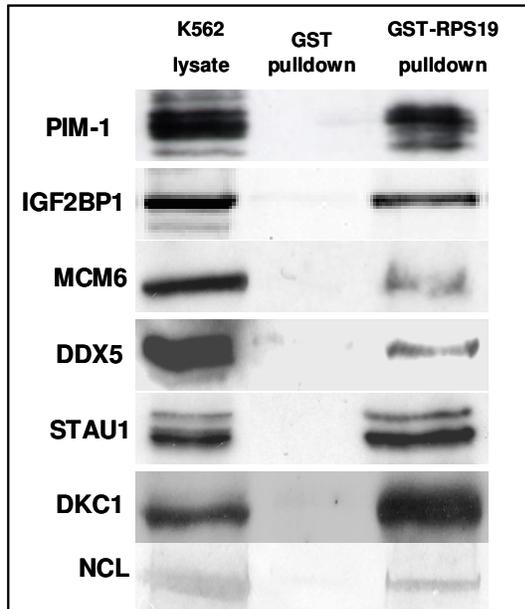
**Figura 23.** Le proteine identificate sono state classificate secondo la loro funzione molecolare (a), localizzazione cellulare (b), processo biologico (c), in accordo con la banca dati Gene Ontology (PubMed [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)).

## 5.2.2 Validazione dei dati di spettrometria di massa mediante *western blotting* e immunoprecipitazione

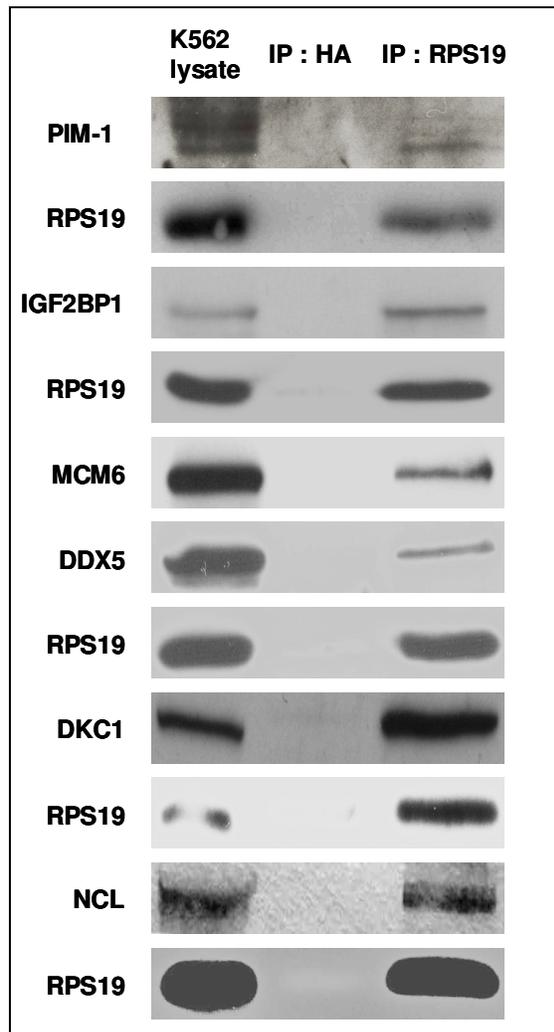
La presenza delle proteine PIM1 (*serine-threonine Kinase*), IGF2BP1 (*insulin-like growth factor 2 binding protein 1*), MCM6 (*minichromosome maintenance deficient protein 6*), DDX5 (*DEAD box polypeptide 5*), STAU1 (*staufen*), DKC1 (*dyskerin*) e NCL (*nucleolin*) nell'eluato della resina GST-Rps19 è stata validata mediante *immuno-blotting* (figura 24). Infine è stato condotto un esperimento di immunoprecipitazione con l'anticorpo monoclonale anti-Rps19 per valutare che le medesime interazioni ricorrono anche *in vivo*. Le proteine immunoprecipitate sono state risolte su un gel d'acrilammide e analizzate per *western blotting* con anticorpi specifici (figura 25). I controlli confermano la specificità d'interazione tra Rps19 e le proteine testate.

## 5.2.3 Confronto tra l'analisi *in silico* e l'analisi sperimentale

L'analisi proteomica *in silico* consente di classificare i ligandi come interattori diretti e indiretti della proteina Rps19. Le banche dati HPRD e PubMed mostrano 4 proteine che interagiscono con Rps19 direttamente (PIM, FGF2, *complement component 5 receptor*, *Rps19 binding protein*) e 32 indirettamente (tabella 13). In questo lavoro sperimentale non sono stati trovati nella lista dei ligandi alcun interattore primario di Rps19. La presenza di PIM nella frazione eluita da GST-Rps19 è stata validata mediante Western Blotting. La scarsa abbondanza non ne ha consentito l'identificazione per spettrometria di massa. Tuttavia è noto dalla letteratura che suddetti interattori interagiscono con Rps19 in condizioni particolari. La proteina FGF2 è riportata interagire con Rps19 libera: l'interazione è stata studiata mediante un esperimento di purificazione GST su estratti proteici citoplasmatici di cellule NIH3T3 [103]. In merito al *complement component 5 receptor* l'interazione subentra quando Rps19 dimerizza in seguito alla formazione di un legame covalente dovuto a transglutaminazione [104]. Nel caso della proteina *Rps19 binding protein* non è disponibile un anticorpo che consenta una validazione dei dati prodotti per spettrometria di massa. L'analisi ha indicato nel nostro lavoro sperimentale la presenza di proteine che interagiscono l'un l'altra: DDX5, PES, DDX21, GTPBP4, NOL5A E NCL interagiscono con RPL4, RPL6, RPL7a, RPL10a, RPLP0 e RPS3. Né le banche dati consultate né dati di letteratura avevano, in precedenza, mostrato relazioni tra le suddette specie e Rps19.



**Figura 24.** Estratti proteici da cellule K562 sono purificati mediante cromatografia d'affinità con resine derivatizzate con GST e GST-RPS19. Gli eluati sono stati analizzati per *western blotting* con gli anticorpi monoclonali anti- PIM-1, anti-IGF2BP1, anti-MCM6, anti-DDX5, anti-STAU1.



**Figura 25.** Estratti proteici da cellule K562 sono immunoprecipitati con gli anticorpi monoclonali anti-Rps19 e anti-HA come controllo negativo e rivelati con gli anticorpi monoclonali anti-PIM1, anti-IGF2BP1, anti-MCM6, anti-DDX5, anti-DKC1, anti-NCL e anti-Rps19 come controllo.

**Tabella 13. Identificazione dei ligandi proteici di Rps19 mediante analisi *in silico*****Interattori Primari di Rps19**

Complement component 5 receptor 1  
 Fibroblast growth factor 2  
 PIM1  
 Rps19 binding protein

**Interattori Secondari di Rps19**

Complement component 5 receptor 1  
     Complement component 5  
     Rps19  
     GNAI2  
     G protein-coupled receptor 77  
 Fibroblast growth factor 2  
     Apoptosis inhibitor 5  
     Protein arginine N-methyltransferase 1  
     FGF receptor 1  
     Rps19  
     CD44  
     Vitronectin  
     Chemokine, CXC motif, ligand 13  
     Glypican 4  
     Translokina  
     \*Casein kinase II, alpha 1  
     \*RPL6  
     FGF receptor 2  
     FGF receptor 4  
     Syndecan 3  
     FGF binding protein 1  
     Perlecan  
     Platelet factor 4  
     Glypican 3  
     Casein kinase 2, alpha 2  
 PIM1  
     NFATC1  
     Sorting nexin 6  
     p100  
     EBNA2 coactivator p100  
     HP1 beta  
     Nuclear mitotic apparatus protein 1  
     HP1 gamma  
     Dynactin 1  
     Dynein  
     CDC 25A  
     Cyclin dependent kinase inhibitor 1A  
     PTP-U2S phosphatase

HSP90A  
SNX6

---

### 5.3 Discussione

Il lavoro svolto rappresenta il primo approccio proteomico globale su larga scala alla definizione dell'interattoma di Rps19. Il confronto tra l'insieme di proteine purificate dall'esca GST e dall'esca GST-Rps19 ha permesso di identificare 159 proteine. *L'analisi in silico* e le informazioni depositate nella banca dati PubMed mostrano che molte delle proteine identificate interagiscono l'un l'altra. Si può ipotizzare che queste ultime con Rps19 partecipano alla formazione di un complesso o di vari complessi multiproteici.

È noto che le proteine ribosomiali sono coinvolte in differenti stadi della biogenesi del ribosoma e/o in distinte fasi traduzionali [105]. In particolare, molte di esse sono elementi fondamentali per processi di: maturazione degli rRNA, assemblaggio proteico, *folding* degli RNA, trasporto di precursori ribosomiali, stabilizzazione delle subunità strutturate, interazione con altri fattori, anch'essi richiesti nelle biogenesi e nella traduzione [106-108]. È stato suggerito il loro coinvolgimento in processi cotraduzionali, come interazioni con fattori proteici di folding all'uscita del ribosoma, traslocazioni co-traduzionali e attività enzimatiche importanti per la funzione ribosomiale, quale ad esempio l'attività elicastica per gli mRNA studiata nei ribosoma batterici.

L'interattoma identificato si compone, in larga parte, di proteine con una localizzazione nucleare e nucleolare, nonostante la purificazione cromatografica sia stata condotta su un estratto proteico totale: non è stata trovata alcuna contaminazione dovuta a proteine citoplasmatiche abbondanti. Questo sottolinea la funzionalità della proteina di fusione GST-Rps19 anche in presenza del tag. Rps19 è parte della subunità ribosomiale 40S e 14 proteine, tra quelle identificate, mostrano detta localizzazione: S2, S3, S4X, S5, S6, S7, S8, S10, S14, S16, S23, S24, S26, SA.

Poiché molte proteine dell'interattoma appartengono alla subunità nucleolare pre-40S e alla 60S (L3, L4, L6, L7, L7a, L8, L9, L10a, L14, L24, L27a) è possibile che molte di esse, tra cui Rps19, partecipino all'assemblaggio già nella fase antecedente la maturazione dell'rRNA, nel pre-ribosoma 90S. Quest'ultima è una struttura dinamica che si avvale di 150 proteine non ribosomiali con un ampio intervallo di attività, quali RNA elicasi, GTPasi, nucleasi, ATPasi, chinasi [109-110]. Sono state trovate 23 delle 31 proteine i cui ortologi, in lievito, appartengono alla subunità 90S.

Rps19 inoltre condivide molti interattori con una peptidil prolil isomerasi, la parvullina, coinvolta nella biogenesi del ribosoma: L3, L4, L6, L7, L7a, L8, L10a, L14, S3, S4X, S6, S8, SA e DDX18 [111].

Poiché l'interattoma di Rps19 si compone di proteine, quali RPS10, RPS26, RPS23, RPS2 coinvolte nel trasporto delle subunità ribosomiale più piccola dal nucleo al citoplasma, e l'esportina XPO, che controlla l'esporto di entrambe le subunità, è possibile ipotizzare anche per Rps19 la medesima funzione. Due recenti studi avvalorano questa ipotesi suggerendo per Rps19 un ruolo nei primi stadi della maturazione degli rRNA ed il suo esporto dal nucleo [88-89]. Inoltre è noto che molte delle proteine purificate partecipano a processi di maturazione dell'rRNA, RNA elicasi, fibrillarina, *component of the box C-D small nucleolar RNAs*, e Nop56.

Appartengono all'interattoma anche componenti del complesso nucleolare delle RNP H/ACA-box, quali la *dyskerin*, NOLA1 e NOLA3. Questo complesso, che include anche NOLA2, è richiesto per la pseudouridilazione specifica degli rRNA nelle prime fasi della biosintesi del ribosoma [112]. Sia la produzione di rRNA 18S che la pseudouridilazione degli rRNA risultano danneggiate in assenza di una delle quattro proteine del complesso.

Un gruppo di ligandi controllano la sintesi proteica, in particolare sono coinvolte in traslocazioni traduzionali, *signal recognition 68* o in regolazioni traduzionali, IGF2BP1 e STAU1. Sono note in letteratura altre proteine ribosomiali, RPL13 e RPL26 [87-88] che

regolano la traduzione di trascritti specifici. La stessa interazione con PIM1, suggerisce per Rps19 un ruolo analogo [86].

Infine questo studio ha portato all'identificazione di proteine che mostrano nuove funzioni: integrine, componenti del proteasoma e chinasi, suggerendo che molto ancora c'è da capire in merito alla funzione di Rps19.

Dunque il lavoro svolto mostra che Rps19 è coinvolta in meccanismi di maturazione dell'RNA cellulare e forse in processi di controllo traduzionale. E' chiaro che i dati forniti mancano della dimensione spazio-temporale. E' tuttavia interessante l'osservazione che interattori diretti ed indiretti di Rps19 sono proteine coinvolte in patologie con fenotipi prossimi alla DBA [90]. In particolare:

- ✓ DKC1 è responsabile di una discheratosi congenita (MIM 305000)
- ✓ RPL24, le cui mutazioni in topo producono ritardo nella crescita e malformazioni scheletriche
- ✓ TCOF1 è responsabile della sindrome Treacher-Collins (MIM 154500), che mostra malformazioni comuni alla DBA e interagisce con NOL5A e UBTF
- ✓ SBDBS (MIM 260400), responsabile della sindrome Schwachman-Diamond, interagisce con la nucleolina.

La definizione dei ligandi molecolari di nota funzione della proteina Rps19 può chiarire le sue attività e far luce sui processi cellulari e/o regolativi in cui la stessa resta coinvolta. Una dettagliata conoscenza della funzione svolta da Rps19 può aiutare a definire il meccanismo d'azione che la associa alla patogenesi DBA e stabilire un collegamento tra malattie ribosomiali e meccanismi patogenetici.

## CONCLUSIONI

In questo lavoro di tesi sono stati studiati ed applicati nuovi protocolli sperimentali volti alla caratterizzazione molecolare di una serie di patologie ancora non dettagliatamente definite. Essendo le proteine i principali effettori delle funzioni vitali di un organismo, l'approccio sperimentale seguito è stato focalizzato sullo studio delle basi molecolari di patologie attraverso l'analisi di corredi proteici. In particolare sono state studiate:

- ✓ Sferocitosi ereditarie, HS
- ✓ Sindrome di DiGeorge, DGS
- ✓ Anemia Diamond Blackfan, DBA

I risultati ottenuti hanno rispettivamente permesso:

- ✓ Rivelare relazioni struttura-funzione anomale in patologie dovute a difetti di membrana. E' stato confrontato il profilo di espressione delle proteine di membrana dei globuli rossi in campioni provenienti da cellule di età differente in soggetti sani e patologici per HS. Studi differenziali hanno mostrato il diverso comportamento di una proteina, l'annexina VII in cellule di età differente sia in soggetti sani che patologici per HS.
- ✓ Individuare i bersagli dell'attività trascrizionale di Tbx1. Sono state identificate circa 50 proteine, la cui espressione è condizionata dall'attività cellulare di Tbx1. Le funzioni biologiche trovate mostrano che molte delle proteine identificate sono coinvolte in processi di nucleotide tri-fosfatasi (NTPasi) e di legame a DNA, RNA e proteine. Infine è stato evidenziato, attraverso la down-regolazione di due proteine Adh2 e PCNA, un collegamento tra l'espressione di Tbx1 e il metabolismo dell'acido retinoico.
- ✓ Definire l'interattoma della proteina ribosomiale Rps19. L'insieme identificato si compone in larga parte di proteine che mostrano una localizzazione nucleare e nucleolare, e funzioni importanti all'assemblaggio delle subunità ribosomiali, facendo ipotizzare per la stessa Rps19 un ruolo nella biogenesi del ribosoma in eucarioti.

**BIBLIOGRAFIA**

- [1]. Csete ME, Doyle JC. Reverse engineering of biological complexity. *Science* 2002, **295** (5560): 1664-9
- [2]. Albert B. The cell as a collection of protein machines: preparing the next generation of Molecular biologists. *Cell* 1998, **92**: 291-294
- [3]. Tyres M, Mann M. From genomics to proteomics. *Nature* 2003, **422**(6928): 193-7
- [4]. White S L, Gharbi S, Bertani M F, Chan H L, Waterfield M D, Timms J F. Cellular responses to ErbB-2 overexpression in human mammary luminal epithelial cells: comparison of mRNA and protein expression. *Br J Cancer*. 2004, **90**(1): 173-81
- [5]. Monti M, Orru S, Pagnozzi D, Pucci P. Interaction proteomica. *Biosci Rep* 2005, **25**(1-2):45-56
- [6]. Gavin AC et al. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 2002, **415**(6868):141-7
- [7]. Lewis TS, Hunt JB, Aveline LD, Jonscher KR, Louie DF, Yeh JM, Nahreini TS, Resing KA, Ahn NG. Identification of novel MAP kinase pathway signaling targets by functional proteomics and mass spectrometry. *Mol Cell* 2000, **6**(6):1343-54
- [8]. Pandey A, Mann M. Proteomics to study genes and genomes. *Nature* 2000, **405**(6788):837-46
- [9]. O' Farrell P H. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 1975, **250**(10): 4007-21
- [10]. Gorg A, Obermaier C, Boguth G, Csordas A, Diaz JJ, Madjar JJ. Very alkaline immobilized pH gradients for two-dimensional electrophoresis of ribosomal and nuclear proteins. *Electrophoresis* 1997, **18**(3-4):328-37
- [11]. Govorun VM, Archakov AI. Proteomic technologies in modern biomedical science. *Biochemistry (Mosc)* 2002, **67**(10):1109-23
- [12]. Patton JL, Lee WA, Pai YC. Relative stability improves with experience in a dynamic standing task. *Exp Brain Res* 2000, **135**(1):117-26
- [13]. Honore B, Ostergaard M, Vorum H. Functional genomics studied by proteomics. *Bioassays* 2004, **26**(8):901-15
- [14]. Hunt SM, Thomas MR, Sebastian LT, Pedersen SK, Harcourt RL, Sloane AJ, Wilkins MR. Optimal replication and the importance of experimental design for gel-based quantitative proteomics. *J Proteome Res* 2005, **4**(3):809-19
- [15]. Gharbi S, Gaffney P, Yang A, Zvelebil MJ, Cramer R, Waterfield MD, Timms JF. Evaluation of two-dimensional differential gel electrophoresis for proteomic expression analysis of a model breast cancer cell system. *Mol Cell Proteomics* 2002, **1**(2): 91-8
- [16]. Patton W F. Detection technologies in proteome analysis. *Journal of Chromatography B* 2002, **771** 3–31

- [17]. Mann M, Hendrickson RC, Pandey A. Analysis of proteins and proteomes by mass spectrometry. *Annu. Rev. Biochem* 2001, **70**: 437-73
- [18]. Aebersold R. A mass spectrometric journey into protein and proteome research. *J Am Soc Mass Spectrom* 2003, **14**(7):685-95
- [19]. Lin D, Tabb DL, Yates JR. Large-scale protein identification using mass spectrometry. *Biochim Biophys Acta* 2003, **1646**(1-2):1-10
- [20]. Aebersold R, Goodlett DR. Mass spectrometry in proteomics. *Chem. Rev.* 2001 **101**: 269-95
- [21]. Mann M, Hendrickson RC, Pandey A. Analysis of proteins and proteomes by mass spectrometry. *Annu. Rev. Biochem.* 2001 **70**: 437-73
- [22]. Cottrell JS. Protein identification by peptide mass fingerprinting. *Pept Res.* 1994, **7**(3):115-24
- [23]. Yates JR. Database searching using mass spectrometry data. *Electrophoresis* 1998, **19**(6):893-900.
- [24]. Godovac-Zimmermann J, Brown LR. Perspective for mass spectrometry and functional proteomics. *Mass Spectrom Rev.* 2001, **20**: 1-57
- [25]. Galletti P, Ingrosso D, Nappi A, Gragnaniello V, Iolascon A, Pinto L. Increased methyl esterification of membrane proteins in aged red-blood cells. Preferential esterification of ankyrin and band-4.1 cytoskeletal proteins. *Eur J Biochem.* 1983, **135**: 25-28
- [26]. Reliene R, Mariani M, Zanella A, Reinhart WH, Ribeiro ML, del Giudice EM, Perrotta S, Iolascon A, Eber S, Lutz HU. Splenectomy prolongs in vivo survival of erythrocytes differently in spectrin/ankyrin- and band 3-deficient hereditary spherocytosis. *Blood.* 2002, **100**: 2208-2215
- [27]. Seow TK, Ong SE, Liang RC, Ren EC, Chan L, Ou K, Chung MC. Twodimensional electrophoresis map of the human hepatocellular carcinoma cell line, HCC-M, and identification of the separated proteins by mass spectrometry. *Electrophoresis* 2000, **21**:1787-1813
- [28]. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 1970, **227**: 680-685
- [29]. Towbin H, Staenelin T, Gordon J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Biotechnology* 1992. **24**: 145-149
- [30]. Iolascon A, D'Agostaro G, Perrotta S, Izzo P, Tavano R, Miraglia del Giudice B. Congenital dyserythropoietic anemia type II: molecular basis and clinical aspects. *Haematologica* 1996, **81**(6):543-59
- [31]. Liu SC, Derick LH, Palek J. Visualization of the hexagonal lattice in the erythrocyte membrane skeleton. *J Cell Biol.* 1987, **104**(3):527-36

- [32]. Delaunay J, Alloisio N, Morle L, Baklouti F, Dalla Venezia N, Maillet P, Wilmotte R. Molecular genetics of hereditary elliptocytosis and hereditary spherocytosis. *AnnGenet.* 1996, **39**:209-21
- [33]. Eber SW, Gonzalez JM, Lux ML, Scarpa AL, Tse WT, Dornwell M, Herbers J, Kugler W, Ozcan R, Pekrun A, Gallagher PG, Schroter W, Forget BG, Lux SE. Ankyrin-1 mutations are a major cause of dominant and recessive hereditary spherocytosis. *Nat Genet.* 1996, **13**:214-218
- [34]. Palek J, Lambert S. Genetics of the red cell membrane skeleton. *Semin Hematol.* 1990, **27**(4):290-332.
- [35]. Delaunay J, Alloisio N, Morle L, Pothier B. The red cell skeleton and its genetic disorders. *Mol Aspects Med.* 1990, **11**(3):161-241
- [36]. Beutler E, Miwa S, Palek J. Hemolytic anemias. *Rev Invest Clin.* 1994 Suppl:162-8
- [37]. Jarolim P, Murray JL, Rubin HL, Taylor WM, Prchal JT, Ballas SK, Snyder L, Chrobak L, Melrose WD, Brabec V, Palek J. Characterization of 13 novel bandgene defects in hereditary spherocytosis with band 3 deficiency. *Blood.* 1996, **88**: 4366-4374
- [38]. Alloisio N, Texier P, Vallier A, Ribeiro M, Morle L, Bozon M, Bursaux E, Maillet P, Goncalves P, Tanner MJ, Tamagnini G, Delaunay J. Modulation of clinical expression and band 3 deficiency in hereditary spherocytosis. *Blood* 1997, **90**: 414-420
- [39]. Yawata Y, Kanzaki A, Inoue T, Ata K, Wada H, Okamoto N, Higo I, Yawata A, Sugihara T, Yamada O. Red cell membrane disorders in the Japanese population: clinical, biochemical, electron microscopic, and genetic studies. *Int J Hematol.* 1994, **60**: 23-38
- [40]. Iolascon A, Miraglia del Giudice E, Perrotta S, Pinto L, Fiorelli G, Cappellini DM, Vasseur C, Bursaux E, Cuttillo S. Hereditary spherocytosis (HS) due to loss of anion exchange transporter. *Haematologica* 1992, **77**:450-456
- [41]. Beutler E (2001). DNA-based diagnosis of red cell enzymopathies: how we threw out the baby with the bathwater. *Blood* 2001, **97**(11):3325
- [42]. Reliene R, Mariani M, Zanella A, Reinhart WH, Ribeiro ML, del Giudice EM, Perrotta S, Iolascon A, Eber S, Lutz HU. Splenectomy prolongs in vivo survival of erythrocytes differently in spectrin/ankyrin- and band 3-deficient hereditary spherocytosis. *Blood.* 2002, **100**: 2208-2215
- [43]. Wallin E, Von Heijne G. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci.* 1998, **7**: 1029-1038
- [44]. Santoni V, Molloy M, Rabilloud T. Membrane proteomics: use of additive main effects with multiplicative interaction model to classify plasma membrane proteins according to their solubility and electrophoretic properties. *Electrophoresis* 2000, **21**(16):3329-44
- [45]. Tastet C, Charmont S, Rabilloud T. Structure-efficiency relationships of zwitterionic detergents as protein solubilizer in two dimensional electrophoresis. *Proteomics* 2003, **3**: 111-121
- [46]. Salzer U, Hinterdorfer P, Hunger U, Borken C, Prohaska R. Ca<sup>++</sup>-dependent vesicle release from erythrocytes involves stomatin-specific lipid rafts, synexin (annexin VII), and sorcin. *Blood* 2002, **99**: 2569-2577

- [47]. Low TY, Seow TK, Chung MC. Separation of human erythrocyte membrane-associated proteins with one-dimensional and two-dimensional gel electrophoresis followed by identification with matrix-assisted laser desorption/ionization-time of flight mass spectrometry. *Proteomics* 2002, **2**: 1229-1239
- [48]. Kakhniashvili DG, Bulla LAJ, Goodman SR. The human erythrocyte proteome: analysis by ion trap mass spectrometry. *Mol Cell Proteomics* 2004, **3**: 501-509
- [49]. Bruschi M, Seppi C, Arena S, Musante L, Santucci L, Balduini C, Scaloni A, Lanciotti M, Righetti PG, Candiano G. Proteomic analysis of erythrocyte membranes by soft immobilized gels combined with differential protein extraction. *J. proteome Res.* 2005, **4**:1304-1309
- [50]. Miraglia del Giudice E, Perrotta S, Lombardi C, Iolascon A. Decision making at the bedside: diagnosis of hereditary spherocytosis in a transfused infant. *Haematologica.* 1998, **83**(4):347-9
- [51]. Mandal M, Mandal A, Das S, Chakraborti T, Chakraborti S. Identification, purification and partial characterization of tissue inhibitor of matrix metalloproteinase-2 in bovine pulmonary artery smooth muscle. *Mol Cell Biochem.* 2003, **254**(1-2):275-87
- [52]. Herr C, Clemen CS, Lehnert G, Kutschkow R, Picker SM, Gathof BS, Zamparelli C, Schleicher M, Noegel A. Function, expression and localization of annexin A7 in platelets and red blood cells: insights derived from an annexin A7 mutant mouse. *BMC Biochemistry.* 2003, **4**:1-11
- [53]. Lindsay EA, Vitelli F, Su H, Morishima M, Huynh T, Pramparo T, Jurecic V, Ogunrinu G, Sutherland HF, Scambler PJ, Bradley A, Baldini A. Tbx1 haploinsufficiency in the DiGeorge syndrome region causes aortic arch defects in mice. *Nature* 2001, **410**(6824):97-101
- [54]. Bulfone A, Smiga S, Shimamura K, Peterson A, Puelles L, Rubenstein J. T-brain-1: a homolog of Brachyury whose expression defines molecularly distinct domains within the cerebral cortex. *Neuron.* 1995, **15**: 63-78
- [55]. Chapman DL, Agulnik I, Hancock S, Silver LM, Papaioannou VE. Tbx6, a mouse T-box gene implicated in paraxial mesoderm formation at gastrulation. *Dev. Biol.* 1996, **125**: 3887-3894
- [56]. Gibson-Brown JJ, Agulnik SI, Chapman DL, Alexiou M, Garvey N, Silver LM, Papaioannou VE. Evidence of role for T-box genes in the evolution of limb morphogenesis and specification of forelimb/hindlimb identity. *Mech. Dev.* 1996, **56**: 93-101
- [57]. Kispert A and Herrmann B G. The Brachyury gene encodes a novel DNA binding protein. *EMBO J.* 1993, **12**: 3211-3220
- [58]. Sinha S, Abraham S, Gronostajski RM, Campbell CE. Differential DNA binding and transcription modulation by three T-box proteins, T, TBX1 and TBX2. *Gene* 2000, **258**(1-2):15-29
- [59]. Naiche LA, Harrelson Z, Kelly RG, Papaioannou VE. T-box genes in vertebrate development. *Annu Rev Genet.* 2005; **39**:219-39
- [60]. Jerome V, Muller R. A synthetic leucine zipper-based dimerization system for combining multiple promoter specificities. *Gene Ther.* 2001, **8**(9):725-9

- [61]. Vitelli F, Lindsay EA, Baldini A. Genetic dissection of the DiGeorge syndrome phenotype. *Cold Spring Harb Symp Quant Biol.* 2002, **67**:327-32
- [62]. Chapman DL, Agulnik I, Hancock S, Silver LM, Papaioannou VE. Tbx6, a mouse T-Box gene implicated in paraxial mesoderm formation at gastrulation. *Dev Biol.* 1996, **180**(2):534-42
- [63]. Garg V. et al. Tbx1, a DiGeorge syndrome candidate gene, is regulated by sonic hedgehog during pharyngeal arch development. *Dev Biol* 2001 **235**: 62–73
- [64]. Vitelli F, Morishima M, Taddei I, Lindsay E A, Baldini A. Tbx1 mutation causes multiple cardiovascular defects and disrupts neural crest and cranial nerve migratory pathways. *Hum Mol Genet* 2002 **11**: 915–922
- [65]. Lindsay EA, Baldini A. Recovery from arterial growth delay reduces penetrance of cardiovascular defects in mice deleted for the DiGeorge syndrome region. *Hum Mol Genet.* 2001 **10**: 997–1002
- [66]. Jerome LA, Papaioannou VE. DiGeorge syndrome phenotype in mice mutant for the T-box gene, Tbx1. *Nature Genetics.* 2001 **27**: 286–291
- [67]. Merscher S et al. TBX1 is responsible for Cardiovascular Defects in Velo-Cardio-Facial/DiGeorge Syndrome. *Cell* 2001 **104**: 619–629
- [68]. Xu H et al. Tbx1 has a dual role in the morphogenesis of the cardiac outflow tract. *Development* 2004 **131**: 3217–27
- [69]. Lin Q, Schwarz J, Bucana C, Olson E N. Control of mouse cardiac morphogenesis and myogenesis by transcription factor MEF2C. *Science* 1997 **276**: 1404–7
- [70]. Lindsay EA. et al. Congenital Heart Disease in Mice Deficient for the DiGeorge Syndrome Region. *Nature* 1999 **401**: 379–383
- [71]. Prescott K, Ivins S, Hubank M, Lindsay EA, Baldini A, Scambler P. Microarray analysis of the Df1 mouse model of the 22q11 Deletion Syndrome. *Hum Genet.* 2005, **116**: 486–496
- [72]. Chen A, Licht JD, Wu Y, Hellinger N, Scher W, Waxman S. Retinoic acid is required for and potentiates differentiation of acute promyelocytic leukemia cells by nonretinoid agents. *Blood* 1994, **84**(7): 2122-9
- [73]. Mark M, Ghyselinck NB, Chambon P. Retinoic acid signalling in the development of branchial arches. *Curr Opin Genet Dev.* 2004, 14(5):591-8
- [74]. Guris DL, Duester G, Papaioannou VE, Imamoto A. Dose-dependent interaction of Tbx1 and Crkl and locally aberrant RA signaling in a model of del22q11 syndrome. *Dev Cell.* 2006, **10**(1): 81-92
- [75]. Kohler T, Prols F, Brand-Saberi B. PCNA in situ hybridization: a novel and reliable tool for detection of dynamic changes in proliferative activity. *Histochem Cell Biol.* 2005, **23**(3):315-27
- [76]. Martin PJ, Lardeux V, Lefebvre P. The proliferating cell nuclear antigen regulates retinoic acid receptor transcriptional activity through direct protein-protein interaction. *Nucleic Acids Res.* 2005, **33**(13): 4311-21

- [77]. Campagnoli MF, Garelli E, Quarello P, Carando A, Varotto S, Nobili B, Longoni D, Pecile V, Zecca M, Dufour C, Ramenghi U e Dianzani I. Molecular basis of Diamond-Blackfan anemia, new findings from the Italian registry and a review of the literature. *Haematologica* 2004, **89**:480-9
- [78]. Draptchinskaia N, Gustavsson P, Andersson B, Pettersson M, Willig T-N, Dianzani I, Ball S, Tchernia G, Klar J, Matsson H, Tentler D, Mohandas N, Carlsson B, Dahl N. The gene encoding ribosomal protein S19 is mutated in Diamond-Blackfan anaemia. *Nat. Genet.* 1999, **21**:169-75
- [79]. Willig TN, Draptchinskaia N, Dianzani I, Ball S, Niemeyer C, Ramenghi U, Orfali K, Gustavsson P, Garelli E, Brusco A, Tiemann C, Perignon JL, Bouchier C, Cicchiello L, Dahl N, Mohandas N, Tchernia G. Mutations in ribosomal protein S19 gene and Diamond Blackfan Anemia, wide variations in phenotypic expression. *Blood* 1999, **94**: 4294-306
- [80]. Da Costa L, Tchernia, G, Gascard P, Lo A, Meerpohl J, Niemeyer C, Chasis JA, Fixler J, Mohandas N. Nucleolar localization of RPS19 protein in normal cells and mislocalization due to mutations in the nucleolar localization signals in 2 Diamond-Blackfan anemia patients, potential insights into pathophysiology. *Blood* 2003, **101**:5039-45
- [81]. Zirwes RF, Kouzmenko AP, Peters JM, Franke WW, Schmidt-Zachmann MS. Topogenesis of a nucleolar protein: determination of molecular segments directing nucleolar association. *Mol Biol Cell.* 1997, **8**: 231-248
- [82]. Russo G, Ricciardelli G, Pietropaolo C. Different domains cooperate to target the human ribosomal L7a protein to the nucleus and to the nucleoli. *J Biol Chem.* 1997, **272**: 5229-5235
- [83]. De Costa L, Naria G, Willing T-N et al. Ribosomal protein S19 (RPS19) expression during erythroid differentiation. *Blood* 2003, **101**: 318-324
- [84]. Léger-Silvestre I, Caffrey JM, Dawaliby R, Alvarez-Arias DA, Gas N, Bertolone SJ, Gleizes PE, Ellis SR. Specific role of yeast homologs of the Diamond-Blackfan anemia associated Rps19 protein in ribosome synthesis. *J. Biol. Chem.* 2005, **280**: 38177-85
- [85]. Ferreira-Cerca S, Poll G, Gleizes PE, Tschochner H, Milkereit P. Roles of eukaryotic ribosomal proteins in maturation and transport of pre-18S rRNA and ribosome function. *Mol. Cell.* 2005, **20**: 263-75
- [86]. Chiocchetti A, Gibello L, Carando A, Aspesi A, Secco P, Garelli E, Loreni F, Angelini M, Biava A, Dahl N, Dianzani U, Ramenghi U, Santoro C, Dianzani I. Interactions between RPS19, mutated in Diamond-Blackfan anemia, and the PIM-1 oncoprotein. *Haematologica* 2005, **90**:1453-1462
- [87]. Mazumder B, Sampath P, Seshadri V, Maitra RK, DiCorleto PE, Fox PL. Regulated release of L13a from the 60S ribosomal subunit as a mechanism of transcript-specific translational control. *Cell* 2003, **115**:187-198
- [88]. Takagi M, Absalon MJ, McLure KG, Kastan MB. Regulation of p53 translation and induction after DNA damage by ribosomal protein L26 and nucleolin. *Cell* 2005, **123**: 49-63
- [89]. Venema J, Tollervey D. Ribosome synthesis in *Saccharomyces cerevisiae*. *Annu. Rev. Genet.* 1999, **33**: 261-311

- [90]. Woolford JLJ, Warner JR. The ribosome and its synthesis. In: Broach JR, Pringle JR, Jones EW (Eds.), *The Molecular and Cellular Biology of the Yeast Saccharomyces: Genome Dynamics, Protein Synthesis and Energetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1991 pp. 587–625
- [91]. Kressler D, Linder P, de La Cruz J. Protein trans-acting factors involved in ribosome biogenesis in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 1999, **19**: 7897–7912
- [92]. Ofengand J, Bachellerie JP. The pseudouridine residues of rRNA: number, location biosynthesis and function. In: Grosjean, H., Benne, R. (Eds.), *Modification and editing of RNA*, ASM Press, Washington 1998, DC, pp. 229–254
- [93]. Bachellerie JP, Cavaille J, Qu LH. Nucleotide modifications of eukaryotic rRNAs: the world of small nucleolar RNA guides revisited. In: Garret R A, Douthwaite S R, Liljas A, Matheson A T, Moore P B, Noller H F (Eds.), *The Ribosome: Structure, Function Antibiotics and Cellular Interactions*, ASM Press, Washington 2000, DC, pp. 191–204
- [94]. Fromont-Racine M, Senger B, Savenau C, Fasiolo F. Ribosome assembly in eukaryotes. *Gene* 2003, **313**: 17-42
- [95]. Nierhaus KH. The assembly of prokaryotic ribosomes. *Biochimie* 1991, **73**: 739–755.
- [96]. Maki JA, Schnobrich DJ, Culver GM. The DnaK chaperone system facilitates 30S ribosomal subunit assembly. *Mol. Cell* 2002 **10**: 129–138
- [97]. Woolford JR. Chaperoning ribosome assembly. *Mol. Cell* 2002 **10**: 8–10
- [98]. Grandi P, Rybin V, Bassler J, et al. 90S pre-ribosomes include the 35S pre-rRNA, the U3 snoRNP, and 40S subunit processing factors but predominantly lack 60S synthesis factors. *Mol. Cell* 2002 **10**: 105–115
- [99]. Nissan TA, Bassler J, Petfalski E, Tollervey D, Hurt E. 60S pre-ribosome formation viewed from assembly in the nucleolus until export to the cytoplasm. *EMBO J.* 2002 **21**: 5539–5547
- [100]. Bassler J, Grandi P, Gadad O et al. Identification of a 60S preribosomal particle that is closely linked to nuclear export. *Mol. Cell* 2001 **8**: 517–529
- [101]. Harnpicharnchai P, Jakovljevic J, Horsey E. Composition and functional characterization of yeast 66S ribosome assembly intermediates. *Mol. Cell* 2001 **8**: 505–515
- [102]. Saveanu C, Namane A, Gleizes PE, Lebreton A, Rousselle JC, Noaillic-Depeyre J, Gas N, Jacquier A, Fromont-Racine M. Sequential protein association to nascent 60S ribosomal particles. *Mol. Cell. Biol.* 2003, **23**(13): 4449-60
- [103]. Soule TF, Al Saati T, Roga S, Amalric F, Bouche G. Fibroblast Growth Factor-2 Interacts with Free Ribosomal Protein S19. *Biochem. Biophys. Res. Commun.* 2001, **289**: 591–596
- [104]. Shibuya Y, Shiokawa M, Nishiura H, Nishimura T, Nishino N, Okabe H, Takagi K, Yamamoto T. Identification of receptor-binding sites of monocyte chemotactic S19 ribosomal protein dimer. *Am. J. Pathol.* 2001, **159**: 2293-2301
- [105]. Blaha G. Structure of the ribosome. In *Protein Synthesis and Ribosome Structure*, Nierhaus, K.H. and Wilson, D.N. eds. (Weinheim, Germany, Wiley-VCH). 2004, pp. 85–106

- 
- [106]. Nierhaus KH. Assembly of the prokaryotic ribosome. In Protein Synthesis and Ribosome Structure, KH. Nierhaus and DN. Wilson, eds. (Weinheim, Germany, Wiley-VCH). 2004pp. 85–106
- [107]. Dresios J, Panopoulos P, Synetos D. Eukaryotic ribosomal proteins lacking a eubacterial counterpart, important players in ribosomal function. *Mol. Microbiol.* 2006, **59**: 1651–1663
- [108]. Kramer G, Rauch T, Rist W, Vorderwulbecke S, Patzelt H, Schulze-Specking A, Ban N, Deuerling E, Bukau B. L23 protein functions as a chaperone docking site on the ribosome. *Nature* 2002, **419**: 171–174
- [109]. Fromont-Racine M, Senger B, Saveanu C, Fasiolo F. Ribosome assembly in eukaryotes. *Gene* 2003, **313**: 42-46
- [110]. Fujiyama S, Yanagida M, Hayano T, Miura Y, Isobe T, Takahashi N. Isolation and proteomic characterization of human Parvulin associating preribosomal ribonucleoprotein complexes. *J. Biol. Chem.* 2002, **277**: 23773–23780
- [111]. Gautier T, Berges T, Tollervey D, Hurt E. Nucleolar KKE/D repeat proteins Nop56p and Nop58p interact with Nop1p and are required for ribosome biogenesis. *Mol. Cell. Biol.* 1997, **17**: 7088–7098
- [112]. Oliver ER, Saunders TL, Tarlé SA, Glaser T. Ribosomal protein L24 defect in Belly spot and tail (Bst), a mouse minute. *Development* 2004, **31**:3907-2

Nel corso di questo progetto di dottorato sono state presentate le seguenti comunicazioni a congresso:

1. Ruoppolo M, Caterino M, Iolascon A, Marrone A, Orrù S, Perrotta S, Savoia M, Salvatore F. Functional proteomics of pathologies related to red cell membrane abnormalities. Sezione Italiana Human Proteome Organization (Primo Congresso Nazionale), Napoli, 26-27/09/2003;
2. Ruoppolo M, Caterino M, Orrù S, Savoia M, Iolascon A. Salvatore F. Erythroid membrane proteomic. Proteine 2004 (XVII Meeting of the Protein Workgroup), Viterbo, 20-22/5/2004
3. Ruoppolo M, Caterino M, Orrù S, Savoia M, Iolascon A. Salvatore F. Erythroid membrane proteomic. Massa 2005 (An International Symposium on Mass Spectrometry), Roma, 28/6/-1/7/2005
4. Orrù S, Caterino M, Ruoppolo M, Savoia M, Perrotta S, Salvatore F, Iolascon A. Erythroid membrane proteomic. HUPO 4th Annual World Congress, Munich, 28/8/-1/9/2005
5. Caterino M, Baldini A, Fulcoli G, Orrù S, Ruoppolo M, Salvatore F. Proteomic analysis to unravel the function of the transcription factor TBX1. Sezione Italiana Human Proteome Organization (Terzo Congresso Nazionale), Lodi, 29/9- 1/10/2005
6. Ruoppolo M, Orrù S, Caterino M, Chiocchetti A, Gibello L, Carando A, Secco P, Biava A, Aspesi A, Dianzani U, Ramenghi U, Loreni F, Santoro C, Dianzani I. A functional proteomic approach to unruevel ribosomal protein S19 protein partners. Sezione Italiana Human Proteome Organization (Terzo Congresso Nazionale), Lodi, 29 /10- 1/10/ 2005

Il lavoro svolto durante questo progetto di dottorato ha prodotto le seguenti pubblicazioni:

1. Silvennoinen L, Myllyharju J, Ruoppolo M, Orrù S, Caterino M, Kivirikko KI, Identification and characterization of structural domains of human ERp57. Association with calreticulin requires several domains. *J Biol Chem.* 2004, **279**:13607-15
2. de Laurentiis A, Caterino M, Orrù S, Ruoppolo M, Tuccillo F, Masullo M, Quinto I, Scala G, Pucci P, Palmieri C, Tassone P, Salvatore F, Venuta S. Partial purification and MALDI-TOF MS analysis of UN1, a tumor antigen membrane glycoprotein. *Int J Biol Macromol.* 2006, **39**(1-3):122-6
3. Caterino M, Ruoppolo M, Orrù S, Savoia M, Perrotta S, Del Vecchio L, Salvatore F, Stewart G W, Iolascon A. Characterization of red cell membrane proteins as a function of red cell density. Annexin VII in different forms of hereditary spherocytosis. *FEBS Letters.* (In press)
4. Orrù S, Aspesi A, Armiraglio M, Caterino M, Loreni F, Ruoppolo M, Santoro C, Dianzani I. Analysis of RPS19's Interactome. *MCP* (Submitted)

Durante il corso di dottorato in Scienze biotecnologiche la dottoressa Caterino ha atteso all'VIII corso di spettrometria di Massa per dottorandi di Ricerca (Siena), e ha trascorso un periodo di formazione nel laboratorio di Proteomica presso l'Istituto di Ricerca di Biologia Molecolare, IRBM, P. Angeletti (Pomezia)