

# Proteomica

## “Classica”

Si occupa della caratterizzazione del **proteoma** di un **determinato** “soggetto”. Confronta contesti diversi (stimoli, stati patologici etc...).

## PTMs

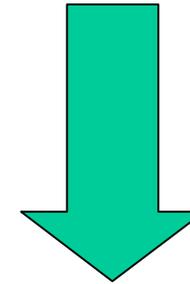
Si occupa dello studio delle variazioni a carico delle modificazioni post-traduzionali delle proteine di un “soggetto”, anche in questo caso **confrontando** contesti diversi (stimoli, stati patologici etc...).

## Interazioni

Si occupa dello studio delle interazioni proteina/proteina e caratterizzazione dei **complessi macromolecolari** (i.e **spliosoma**, **complesso trascrizionale di base**, etc...).

## Struttura

Si occupa della determinazione della struttura delle proteine con approcci a larga scala.



From: **Structural proteomics: developments in structure-to-function predictions - Trends in Biotechnology - 20, 2002**

Over the past few years, proteomics has gained much attention for the high throughput (HT) identification and analysis of expressed proteins from cells – the proteome. Particular interest has been given to the study of expression levels as a consequence of developmental stage or external stimuli such as drugs. Proteomics methods are used to generate a more ‘detailed’ knowledge than that obtained by analysis of DNA sequences because it is possible to identify post-translational modifications and interaction components that modulate the activity of the target. Structural proteomics emerged as a consequence of simultaneous development of HT methodologies and technologies that enabled novel data to be generated with efficiency and speed. Key factors for protein production include HT cloning and expression from multiple vectors in multiple hosts, core domain identification using proteolysis methods and use of expression and detection tags. Protein crystallography made significant improvements over the past decade when freezing crystals at the temperature of liquid nitrogen (cryofreezing), multiple-wavelength anomalous dispersion (MAD) phasing, robotization, automated data collection and the use of synchrotron beamlines were adopted as standards. The improvements in structure determination by multidimensional-NMR using isotope-enriched protein samples include the use of high-field spectroscopy, cryogenic probes, automated spectra assignment and structure determination. The advances in NMR and crystallography methods have made structure determination relatively straightforward if the protein sample has appropriate physical properties, such as solubility, aggregation and homogeneity.

**Scopo della proteomica: fornire informazioni che possono essere di tipo**

## **Differenziale**

**Comprensione di quali sono le variazioni a livello dei componenti proteici coinvolti in un determinato processo cellulare (vale anche per p/p interactions and PTMs)**

**Analisi**

**Comparativa**

## **Totale**

**Comprensione di quali sono i costituenti proteici di una determinata cellula, organello, complesso macromolecolare etc. (vale anche per p/p interactions and PTMs)**

**Analisi su un**

**singolo sistema**

Per capire cosa la proteomica ci può aiutare a comprendere è utile porsi questa domanda:

**Quali sono le domande alle quali la Proteomica può dare una risposta?**

**Quali proteine sono espresse in un determinato tipo cellulare/tessuto/organo/organismo**

**HUPO – Human Proteome Organisation**

- **Human Liver Proteome Project**
- **Plasma Proteome Project**
- **Human Brain Proteome Project**

Goal:

Define and decipher normal brain proteome including polymorphisms, modifications, histological localization as well as identification of brain derived proteins in bodily fluids.

## HUPO Plasma Proteome Project.

The pilot phase of the Plasma Proteome Project (PPP) from 2002 to 2005 was a collaborative international comparative analysis of the experimental approaches and protein identifications from the participating laboratories. As a prerequisite to the comparative analysis, all data were collected and stored in a central sql server database, here at the University of Michigan. Phase 2 of the Human Plasma Proteome Project continues .

**The following links lead to tab delimited text files containing the data created by different queries:**

**1)All Identifications:** All protein identifications from MS/MS experiments submitted by participating laboratories **with either low or high confidence**. There are 9504 proteins in this list.

**2)Confirmed Identifications:** All protein identifications from MS/MS experiments submitted by participating laboratories **with either low or high confidence**, but **identified with two or more distinct peptides**. There are 3020 proteins in this list.

**3)List of 889 high confidence HPPP proteins:** A rigorous statistical approach devised to take into account the length of coding regions in genes and multiple hypothesis testing yielded a reduced set of 889 **proteins with at least 95% confidence** in protein identification.

**3 - 2011 (about 10 years later):1929 proteins at 1% confidence level**

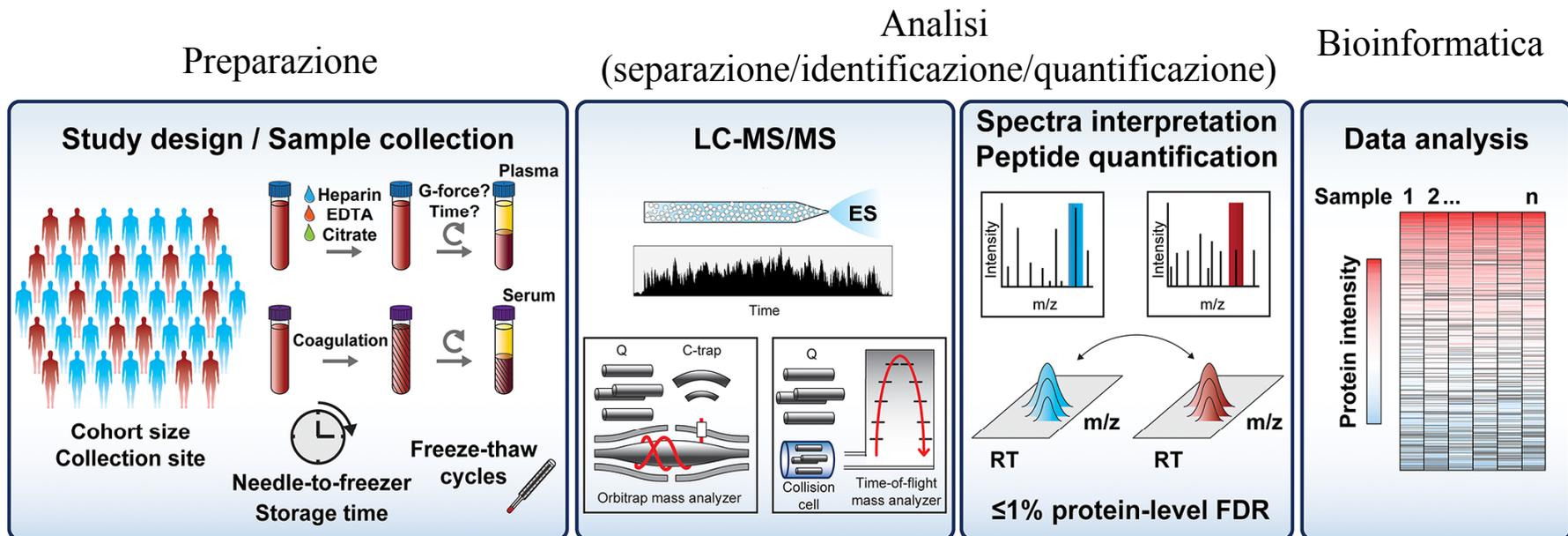
**Concetto di Confidenza nell'identificazione di una proteina – probabilità che il processo d'identificazione sia corretto, ovvero che l'identità sia proprio quella.**

**STATISTICA è fondamentale per una corretta interpretazione dei dati**

HUMAN Proteome: strong protein-level evidence (i.e., PE1 data) for **17 694** proteins

2011 - 1921 (FDR 1%) => 2017 - 3500 (FDR 1%, 20% of human proteome)

FDR: false discovery rate - Percentuale di falsi positivi accettata



J. Proteome Res. 2019, 18, 4085–4097

Domanda biologica

Grandezza del numero di campioni => dati statisticamente significativi

NB: standardizzazione delle operazioni => riproducibilità dei risultati => attendibilità degli stessi (GLP)

Scelta delle piattaforme tecnologiche

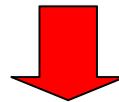
Comparazione dati tra laboratori diversi

# Quali sono le domande alle quali la Proteomica può dare una risposta

**Quali proteine sono differenzialmente espresse oppure modificate (sia in termini qualitativi che quantitativi) confrontando due (o più) fenotipi cellulari**

Un esempio di diversi fenotipi cellulari potrebbe essere:

- Cellule di origine diversa;
- Cellule dello stesso tipo sottoposte a stimoli diversi;
- Cellule dello stesso tipo soggette o meno a condizioni patologiche.



IDENTIFICARE PROTEINE CARATTERISTICHE di UNA  
DETERMINATA CONDIZIONE CELLULARE  
MARKERS



FUNZIONE

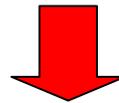
## Quali sono le domande alle quali la Proteomica può dare una risposta

**Quali proteine sono differenzialmete soggette ad una particolare PTM confrontando  
fentotipi cellulari diversi**

Fosfo-proteoma  
Metil-proteoma  
Acetil-proteoma  
Ubiquitin-proteoma

...

Stabilire l'insieme di tutte le proteine soggette alla medesima  
modificazione post-traduzionale e seguirne i cambiamenti in risposta a  
determinate condizione

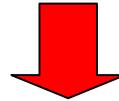


Identificare pathway specificamente attivate in seguito ad uno stimolo  
tramite l'identificazione dei substrati

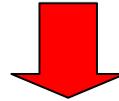
# Quali sono le domande alle quali la Proteomica può dare una risposta

**Quali proteine costituiscono/sono presenti in un determinato complesso macromolecolare/organello**

IDENTIFICAZIONE



RUOLO nel COMPLESSO/ORGANELLO

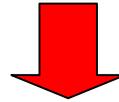


MODULAZIONE ATTIVITA' COMPLESSO/ORGANELLO  
Attraverso PROTEINA

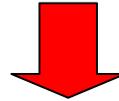
# Quali sono le domande alle quali la Proteomica può dare una risposta

**Quali proteine interagiscono con una determinata proteina (X)**  
**Determinazione del network molecolare di una proteina (X)**

IDENTIFICAZIONE



FUNZIONE dell'INTERAZIONE



MODULAZIONE ATTIVITA della PROTEINA (X)

FUNZIONE della PROTEINA (X)

# Approccio proteomico – GENERALE

(i quattro step principali per lo sviluppo di un approccio proteomico)

1

Ottimizzazione  
Estrazione proteine  
-  
(Preparazione del Campione)

Riuscire a **recuperare** le  
proteine dalla matrice biologica  
d'interesse

2

Selezione – Ottimizzazione  
metodo separativo

Utilizzare dei metodi separativi  
per cercare di **separare** una  
dall'altra le proteine oggetto  
d'indagine

3

Scelta metodo  
Rilevazione

Utilizzare dei metodi che mi  
permettano di “**vedere**” le  
proteine oggetto d'indagine

4

Scelta metodo  
Identificazione

Utilizzare dei metodi che mi  
permettano di ottenere **l'identità**  
delle proteine oggetto d'indagine

# Preparazione del Campione

## Requisiti fondamentali di un metodo per l'estrazione/preparazione del campione:

### ⇒ 1) Efficienza

L'efficienza del metodo per estrarre le proteine dal materiale di partenza deve essere la più elevata possibile, ovvero devo poter ottenere il maggior numero di proteine.

### ⇒ 2) Evitare contaminazioni (= Specificità)

L'omogeneità del campione è importante per non avere nelle proprie analisi proteine che non derivano dal campione d'interesse - problema che si riscontra soprattutto con tessuti (popolazioni cellulari eterogenee)

⇒ **necessaria procedura altamente selettiva per il campionamento (elevata specificità)**

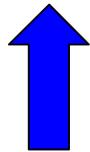
### ⇒ 3) Evitare modificazioni

Processamento del campione lungo (più passaggi) equivale a dire perdita di campione e modificazioni a carico delle proteine ⇒ preferibile adottare **metodi estrattivi con pochi passaggi, rapidi e mantenere il campione a temperature ragionevolmente basse** (attorno a 0 °C) - compatibilmente con i reagenti utilizzati per l'estrazione (alcune sostanze hanno una bassa solubilità a basse temperature).

**ATTENZIONE:** Adottare **mezzi estrattivi compatibili con le analisi che seguiranno** (presenza di sali, detergenti, solventi organici etc. deve essere valutata in funzione del tipo di metodica analitica che si andrà ad adottare). Questo è un punto da tenere in considerazione sempre quando si eseguono determinate operazioni nel contesto di un protocollo scientifico.

# Sistemi biologici

**Proteine**

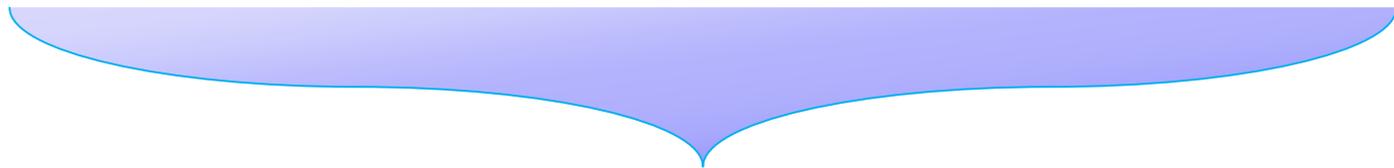


Lipidi

Carboidrati

Acidi nucleici

Co-fattori



VISTI COME INTERFERENTI NELLE ANALISI DI TIPO PROTEOMICO  
I metodi d'estrazione delle proteine devono tener conto di questo aspetto

## Estrazione delle proteine

Isolare le proteine

– distruggere le cellule/tessuti in cui esse sono contenute in modo da renderle solubili -

### Tamponi di estrazione

- Tampone (pH)
- Inibitori di proteasi/altri enzimi
- Detergenti
- Agenti/solventi solubilizzanti



### Metodi di lisi

# Tamponi d'estrazione

Prima considerazione: Condizioni Native vs. Denaturanti

**Condizioni native:** preservare struttura/attività delle proteine d'interesse

NB: esempio nel caso di isolamento di complessi macromolecolari, o qualora si vogliano ottenere informazioni strutturali, etc...

**Condizioni denaturanti:** non è necessario preservare struttura/attività delle proteine d'interesse

NB: esempio nel caso di isolamento di proteine per effettuare analisi differenziali.

**Buffer** (controllo pH dell'estratto) – estremamente importante la dove si voglia preservare l'attività di una determinata proteina. (Non volatili: Glicina [2.6-3.6], Citrato [3-6.2], Acetato [3.6-5.6], Tris [7.2-9.0], Fosfato [5.7-8], Borato [7.6-9.2], etc... – eventualmente a pH con HCl e NaOH). Volatili: Acido Acetico/Acido Formico, Ammonio Carbonato, Ammonio Bicarbonato, etc... – tutti i composti con i quali vengono formati questi tamponi sono volatili. NB: attenzione ad eventuali sali introdotti nella soluzione portando a pH il tampone.

**Inibitori di proteasi:** controllo dell'attività proteolitica naturalmente presente nei campioni

Esistono diversi tipi di proteasi (Serin-, cistein-, aspartico-, e metallo-proteasi), ciascuna delle quali viene inibita da specifici inibitori => cocktail preformati – commercializzati da diverse ditte di prodotti biotecnologici.

**Inibitori di altre attività enzimatiche** (chinasi, fosfatasi, etc).

**Detergenti:** La maggior parte delle proteine sono solitamente solubili in acqua. Ci sono però casi dove questo non si verifica, come ad esempio le proteine di membrana. In altri casi le proteine sono presenti sotto forma di complessi macromolecolari ed è necessario andare ad indagare i singoli costituenti dei vari complessi. I detergenti, avendo una struttura anfipatica (porzione idrofobica e porzione idrofilica) aiutano la solubilizzazione delle proteine. Inoltre possono disturbare i doppi strati lipidici andando a favorire il rilascio delle proteine.

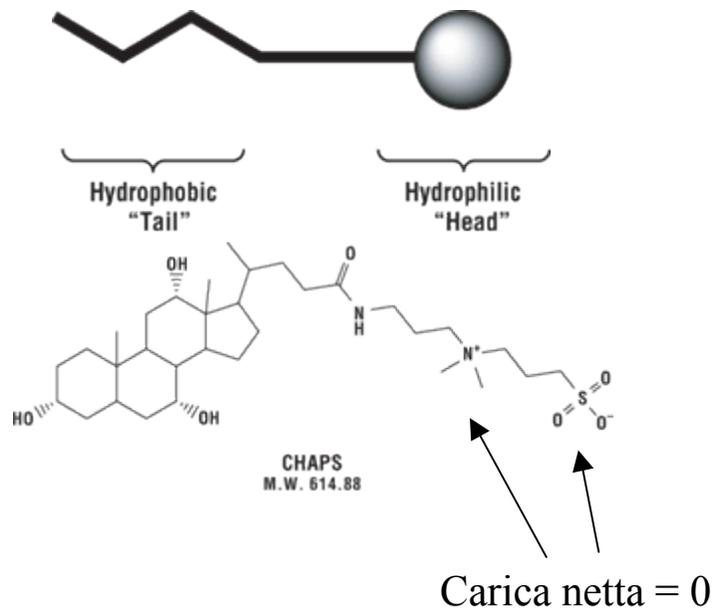
=> detergenti ionici – anionici o cationici – solitamente sono denaturanti

=> Neutri

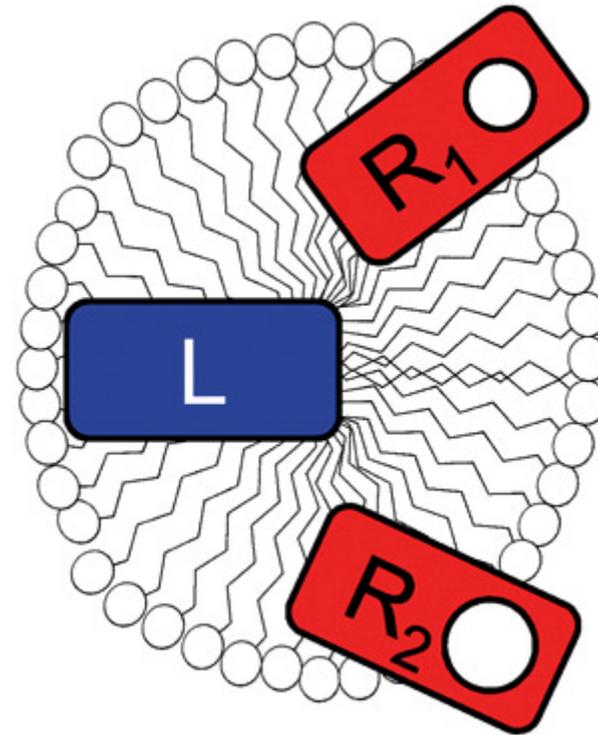
=> zwitterionici (doppia carica, sia positiva che negativa)

**Altri agenti solubilizzanti:** Urea, tiourea, etc - agenti caotropici che aiutano il mantenimento in soluzione di proteine idrofobiche

## Struttura detergenti



## Formazione micelle



NB: la formazione di micelle può comportare il fatto che le molecole siano inserite in strutture di dimensioni specifiche (micelle hanno una dimensione specifica). Questo fatto è da tenere in considerazione quando si effettuano dei trattamenti con sistemi che prevedono un qualche processamento con sistemi che si basano sulla dimensione molecolare (i.e. vedasi gel filtration, ultrafiltrazione, etc)

**Metodi di lisi** – come portare in soluzione le proteine facendole “uscire” dal loro contesto cellulare

Cosa bisogna considerare: procarioti, lieviti, cellule animali e cellule vegetali hanno tutti pareti/membrane cellulari diverse e che sono quindi diversamente resistenti all’azione di detergenti/sistemi meccanici per la rottura di queste “strutture contenitive”.

## Metodi di lisi “SOFT”:

**1)Lisi osmotica:** utilizzo di soluzioni ipotoniche che causano rigonfiamento ed esplosione della cellula con conseguente rilascio delle proteine – non adatto a procarioti con parete cellulare.

**2)Congelamento/scongelamento:** formazioni di cristalli provoca la rottura della parete e della membrana cellulare.

**3)Lisi con detergenti:** detergenti sono in grado di rompere la struttura delle membrana plasmatica facendo uscire le proteine. Attenzione al tipo di detergenti utilizzati.

**4)Lisi enzimatica:** cellule con pareti cellulari possono essere lisate grazie all’utilizzo di enzimi in grado di degradare la parete cellulare stessa. Classico esempio è la parete cellulare (composta da peptidoglicano) dei batteri che può essere degradata ad opera dell’enzima **lisozima** e quella delle cellule vegetali che può essere degradata dalla **cellulasi**

Nota bene: anche se la composizione del mezzo estrattivo non è indicata –  
essa deve essere compatibile con successivi passaggi!

## Metodi di lisi “più forti”:

**1)Sonicazione:** onde ultrasoniche PROVOCANO IL FENOMENO DELLA CAVITAZIONE (formazione di bolle che poi implodono portando a repentini cambiamenti nella temperatura e pressione). Questo favorisce la rottura di interazioni intermolecolari e soprattutto la rottura delle membrane cellulari. Attenzione alla generazione di calore.

**2)Rottura meccanica** tramite pestello

3)Rottura attraverso il **passaggio forzato per un orifizio** molto piccolo

**4)Utilizzo di omogeneizzatore** (attenzione al calore)

**5)Utilizzo di forze abrasive** – sfere di vetro sottoposte a forte agitazione (metodo utilizzato per lieviti)

Nota bene: anche se la composizione del mezzo estrattivo non è indicata – essa deve essere compatibile con successivi passaggi!

## Metodi di estrazione differenti sono spesso **combinati** assieme

### Esempio 1: Estrazione di proteine da cellule eucariotiche

Cellule lavate in PBS e congelate a  $-80^{\circ}\text{C}$  (lisi attraverso congelamento - 1° metodo).

Cellule scongelate e risospese in Urea 7M, Tiourea 2M, CHAPS 4%, etc (Estrazione attraverso utilizzo di agenti caotropici e detergenti – 2° metodo).

Cellule sonicate  $3 \times 10''$  (Estrazione tramite sonicazione – 3° metodo).

### Esempio 2: Estrazione di proteine nucleari

Cellule lavate in PBS e sottoposte a lisi in ambiente iipototonico (Lisi membrana cellulare ma nuclei recuperabili dopo centrifugazione – 1° metodo estrattivo – prefrazionamento nucleo/citoplasma)

Proteine estratte dai nuclei attraverso alta forza ionica (2° metodo estrattivo – prefrazionamento Proteine istoniche – proteine non istoniche).