

# Máster Universitario en Bioinformática

## Proteómica y Bioinformática Estructural

Curso académico 2024-2025



**Universidad**  
Internacional  
de Valencia

**Dra. Magdalena Nikolaeva Koleva**

[magdalena.nikolaeva@professor.universidadviu.com](mailto:magdalena.nikolaeva@professor.universidadviu.com)

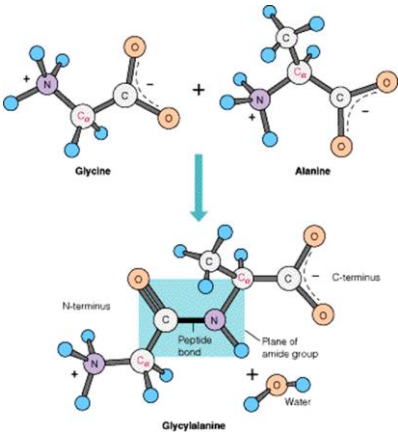
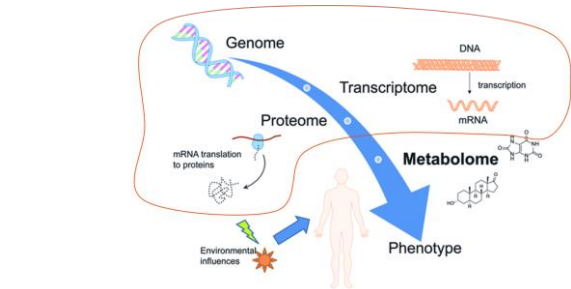
19/09/2024

De:

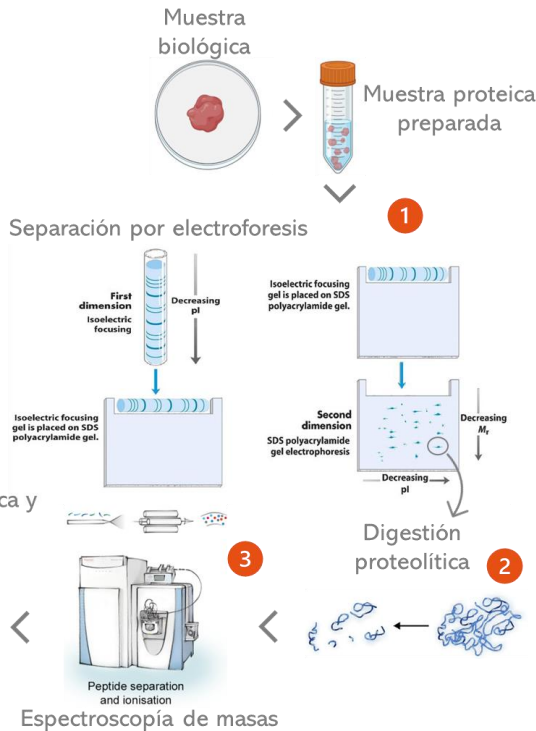
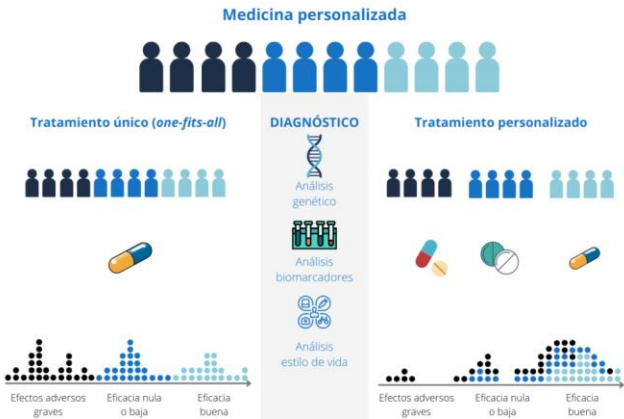
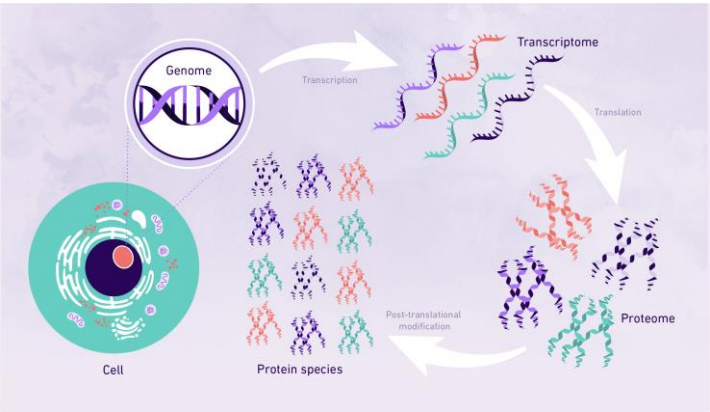


Planeta Formación y Universidades

¿Qué vimos en la anterior sesión?



Aminoácido	Tipo	Abreviatura	Letra	pK <sub>a1</sub> (-COOH)	pK <sub>a2</sub> (-NH <sub>2</sub> )	pK <sub>aR</sub> (R)	pI
Glicina	Neutros apolares	GLI	G	2,34	9,78		6,06
Alanina		ALA	A	2,35	9,69		6,02
Valina		VAL	V	2,32	9,62		5,97
Leucina		LEU	L	2,36	9,64		6,00
Isoleucina		ILE	I	2,36	9,68		6,02
Metionina		MET	M	2,28	9,21		5,75
Prolina	Neutros aromáticos	PRO	P	1,99	10,60		6,30
Fenilalanina		PHE	F	1,83	9,29		5,53
Tirosina		TRY	Y	2,20	9,11	10,07	5,65
Triptófano		TRP	W	2,38	9,39		5,89
Serina	Neutros polares	SER	S	2,21	9,15		5,68
Cisteína		CYS	C	1,96	10,28	8,18	5,07
Treonina		TRE	T	2,71	9,62		6,16
Asparagina		ASG	N	2,02	8,80		5,41
Glutamina	Ácidos	GLN	Q	2,17	9,13		5,65
Ácido aspártico		ASP	D	2,09	9,82	3,86	2,97
Ácido glutámico	Básicos	GLU	E	2,19	9,67	4,25	3,22
Lisina		LYS	K	2,18	8,95	10,53	9,74
Arginina		ARG	R	2,17	9,04	12,48	10,76
Histidina		HIS	H	1,82	9,17	6,00	7,58



# Sesión 2

## Caracterización de expresión proteómica



**Universidad**  
Internacional  
de Valencia

# Temario - Contenidos

Tema 2. Métodos en proteómica: caracterización de la expresión de proteínas

2.1. Separación o fraccionamiento de proteínas

2.2. Caracterización e identificación de proteínas

2.3. Cuantificación y expresión diferencial de proteínas

2.4. Análisis computacional de datos en proteómica

# Temario - Contenidos

Tema 2. Métodos en proteómica: caracterización de la expresión de proteínas

2.1. Separación o fraccionamiento de proteínas → Electroforesis, cromatografía

2.2. Caracterización e identificación de proteínas → Espectrometría de masas

2.3. Cuantificación y expresión diferencial de proteínas

2.4. Análisis computacional de datos en proteómica



Combinación  
cromatografía y  
espectrometría de masas

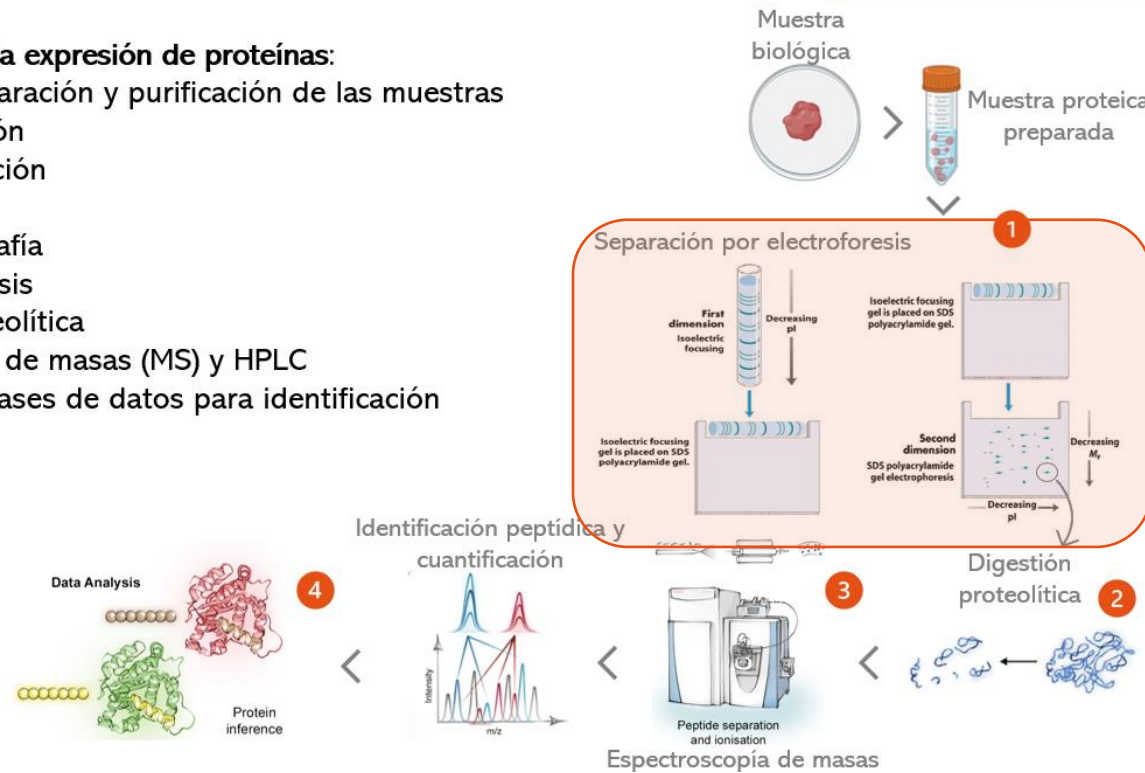
# 2.1. Separación o fraccionamiento de proteínas

## Electroforesis:

### 1.4. Proteómica - metodologías

Caracterización de la expresión de proteínas:

1. Extracción, separación y purificación de las muestras
  - Precipitación
  - Centrifugación
  - Filtración
  - Cromatografía
  - Electroforesis
2. Digestión proteolítica
3. Espectrometría de masas (MS) y HPLC
4. Búsqueda en bases de datos para identificación

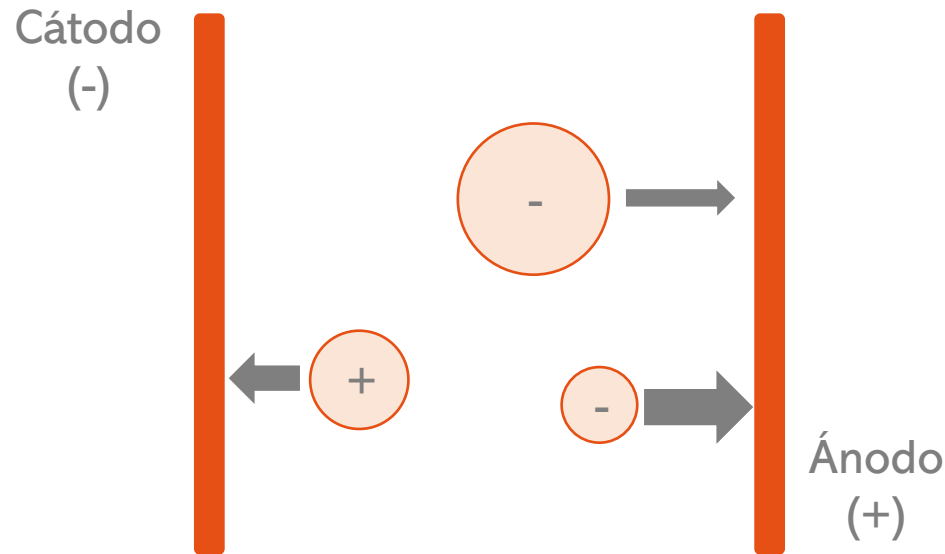


## 2.1. Separación o fraccionamiento de proteínas

### Electroforesis:

Técnica de separación basada en la diferente movilidad de las sustancias en solución bajo la acción de un campo eléctrico.

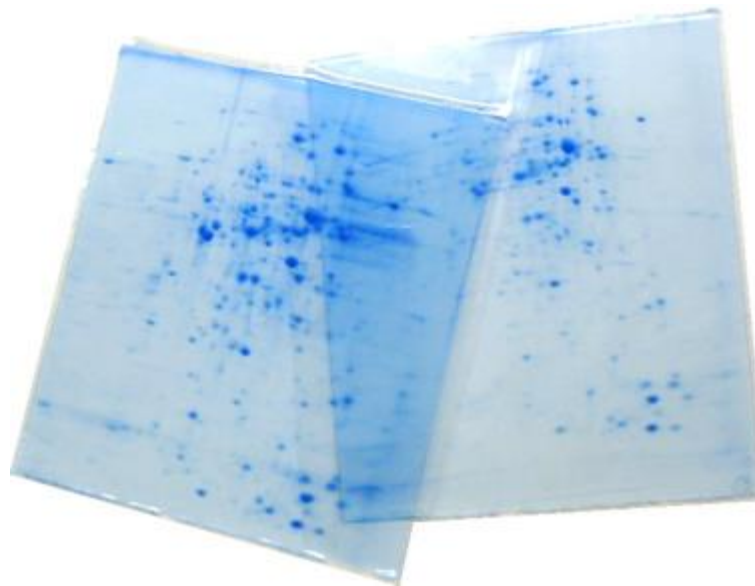
La separación depende de una propiedad específica para cada sustancia: **movilidad electroforética**



## 2.1. Separación o fraccionamiento de proteínas

### Electroforesis bidimensional:

- Primera dimensión: **isoelectroenfoque**
- Segunda dimensión: **SDS-PAGE**



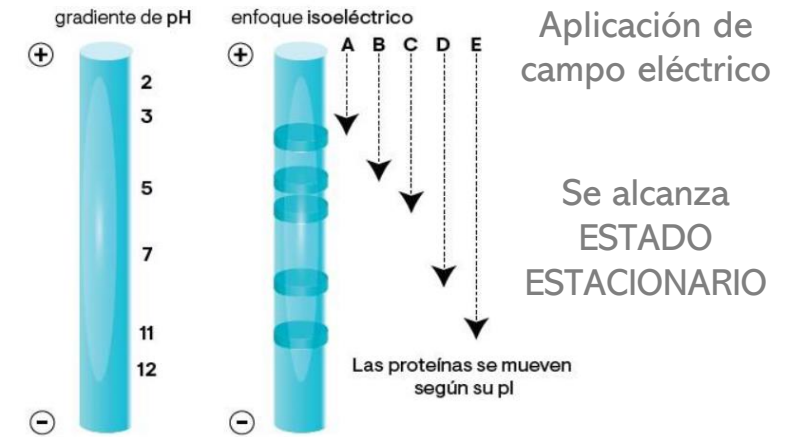
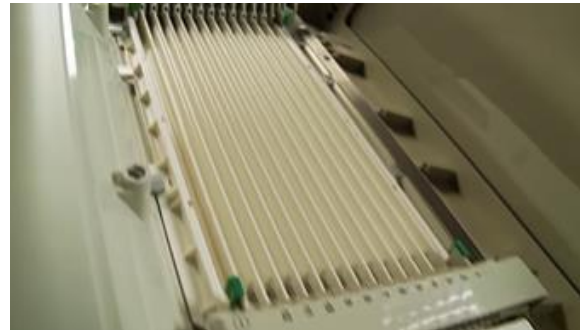
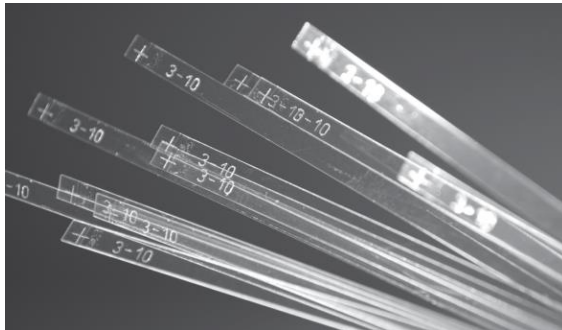
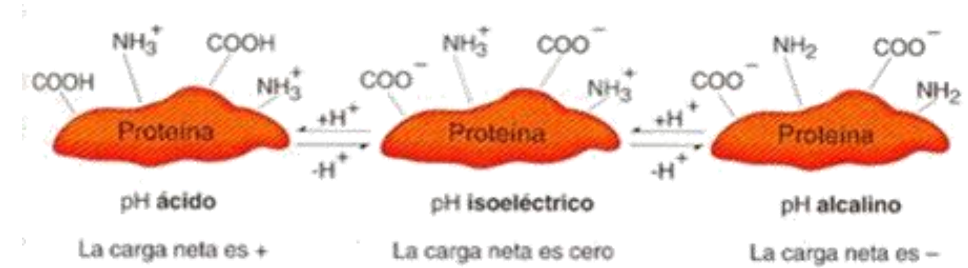


# 2.1. Separación o fraccionamiento de proteínas

## Electroforesis bidimensional:

○ Primera dimensión: **isoelectroenfoque**  
Separación de proteínas en función de su **punto isoeléctrico**.

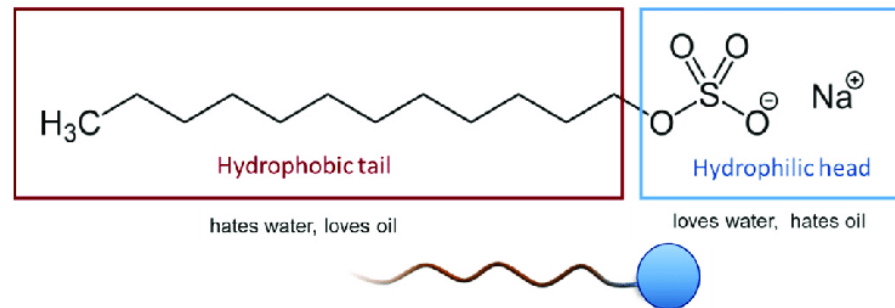
- **Clásico:** gradiente de pH creado **por anfolitos transportadores**. Inestable a tiempos largos. Las proteínas son un anfolito más. Menor reproducibilidad.
- **Moderno:** gradiente de pH **inmovilizado sobre acrilamida (inmovilinas)**.



## 2.1. Separación o fraccionamiento de proteínas

### Electroforesis bidimensional:

- Segunda dimensión: **SDS-PAGE**
  - **Geles de poliacrilamida (PAGE) – electroforesis nativa**
    - Químicamente inertes
    - Transparentes
    - Estables en rango amplio de pH, temperatura y fuerza iónica
    - Fáciles de generar por polimerización de acrilamida + **bis-acrilamida**. El tamaño del poro depende del % de bis-acrilamida vs acrilamida.
    - La polimerización se inicia en presencia de un **iniciador** (TEMED) y un **catalizador** (persulfato amónico, APS).
  - Geles con **SDS (sodium dodecyl sulfate)** – **desnaturalizar** proteínas



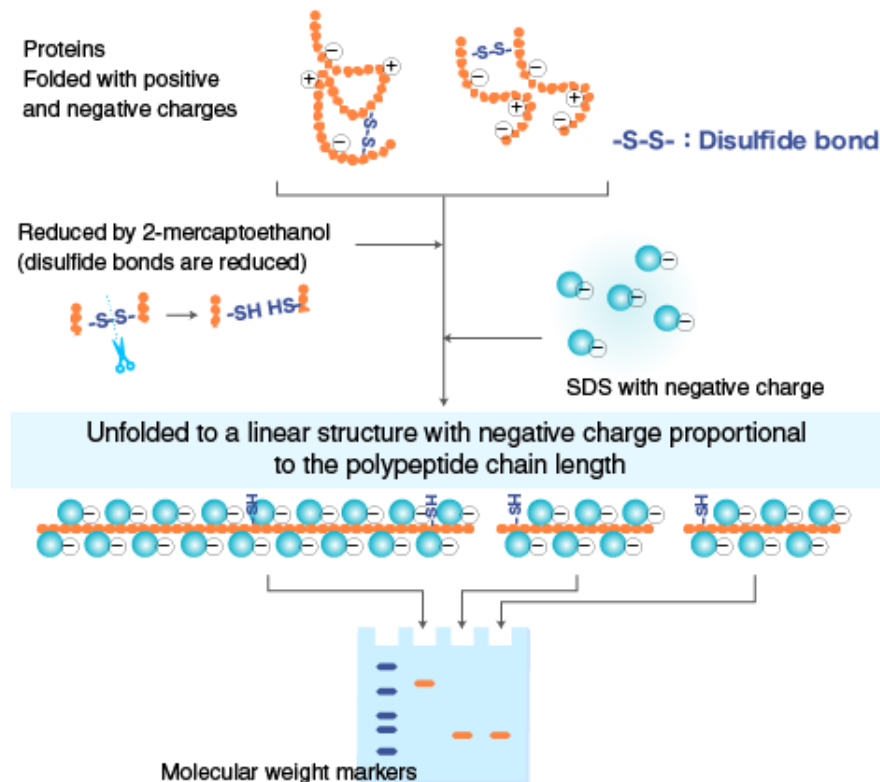
## 2.1. Separación o fraccionamiento de proteínas

### Electroforesis bidimensional:

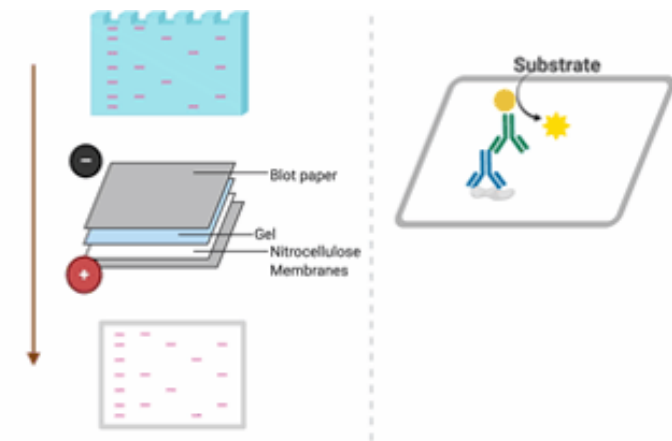
- Segunda dimensión: **SDS-PAGE**

- Geles con **SDS** (sodium dodecyl sulfate) – **desnaturalizar** proteínas

Elimina la influencia de la estructura y la carga y las proteínas se separan únicamente en función de la longitud de su cadena polipeptídica.



**Western blot: electrotransferencia** a otros soportes (nitrocelulosa) para **TÉCNICAS DE INMUNODETECCIÓN**



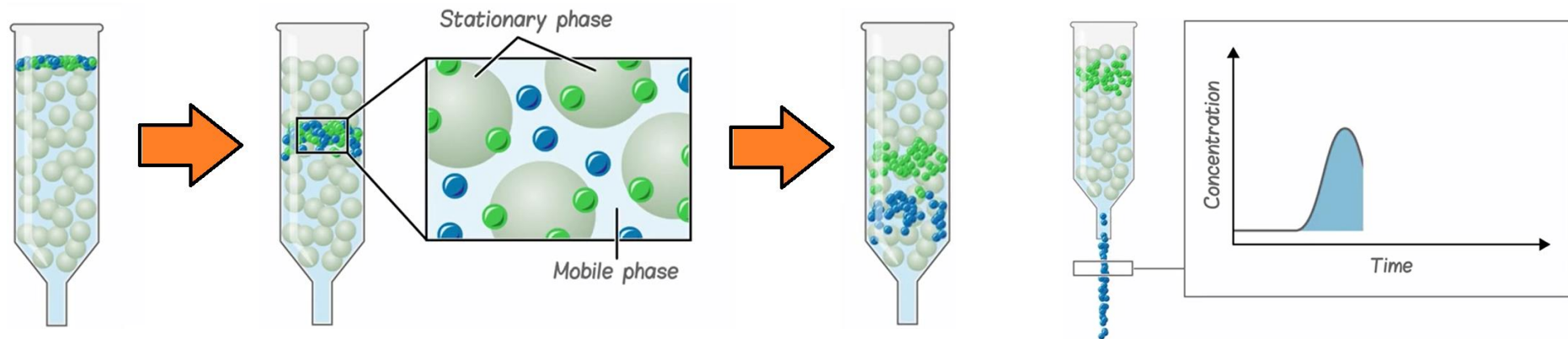
# 2.1. Separación o fraccionamiento de proteínas

## Cromatografía:

Proceso físico de separación en el que los componentes de una mezcla se separan entre una **fase estacionaria** y una **fase móvil**. Ocurre como resultado de diversas etapas de **adsorción-desorción** (elución).

Tipos:

- Cromatografía de afinidad
- Cromatografía de intercambio iónico
- Cromatografía de interacción hidrofóbica
- Cromatografía de exclusión molecular

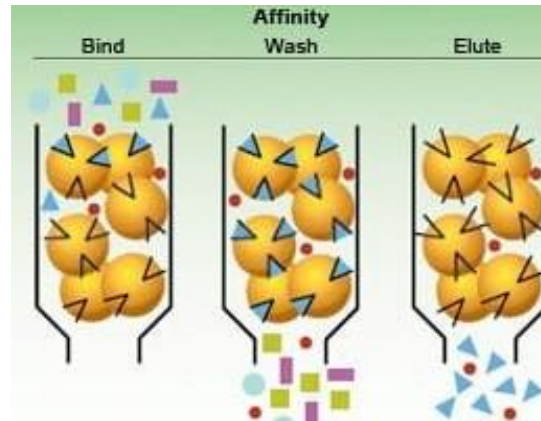


## 2.1. Separación o fraccionamiento de proteínas

### Cromatografía:

Tipos:

- **Cromatografía de afinidad** – **interacción específica** entre la matriz de la columna cromatográfica y una o un grupo de proteínas de la muestra.

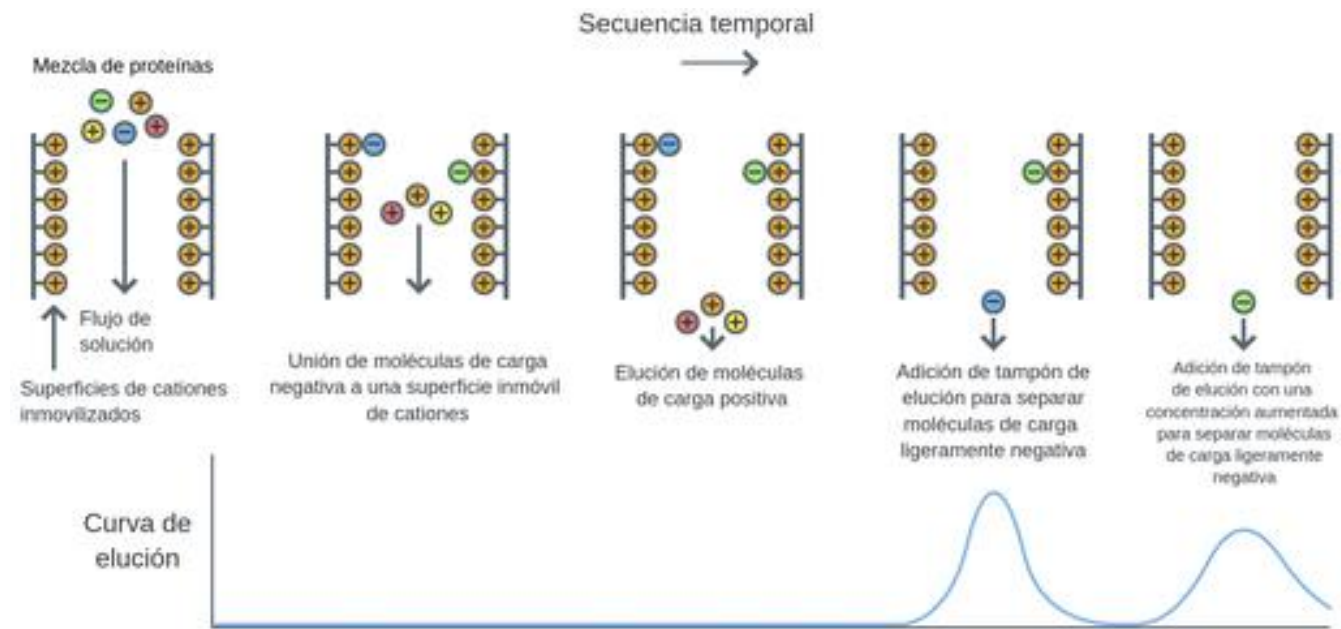


## 2.1. Separación o fraccionamiento de proteínas

### Cromatografía:

Tipos:

- **Cromatografía de intercambio iónico** – los componentes de la muestra se separan según su **carga**. Selección apropiada de **fase estacionaria** y el **tampón de carga** (pH). Para la elución normalmente se usa un **gradiente**.

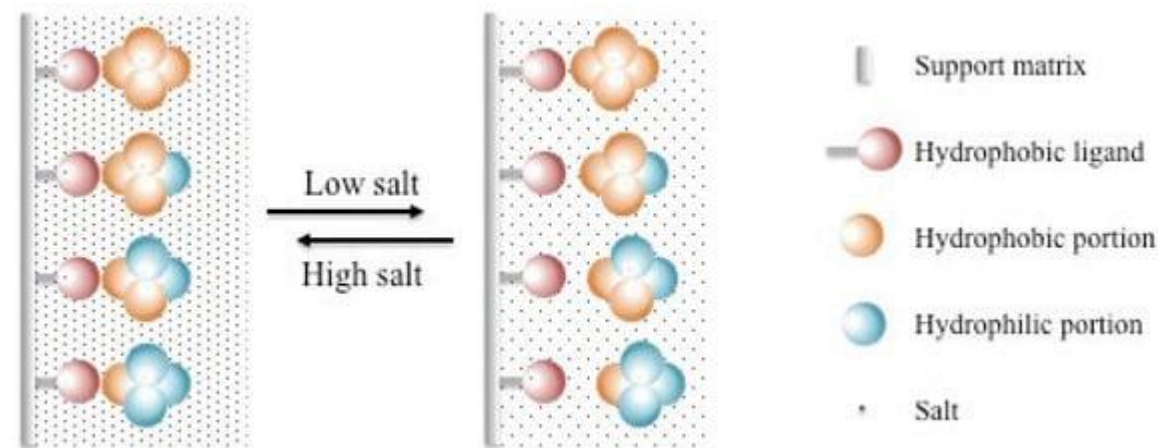


## 2.1. Separación o fraccionamiento de proteínas

### Cromatografía:

Tipos:

- **Cromatografía de interacción hidrofóbica** – separa las moléculas en función de su **hidrofobicidad**. La fase estacionaria presenta grupos hidrofóbicos. Para la elución se aplica un gradiente salino decreciente.



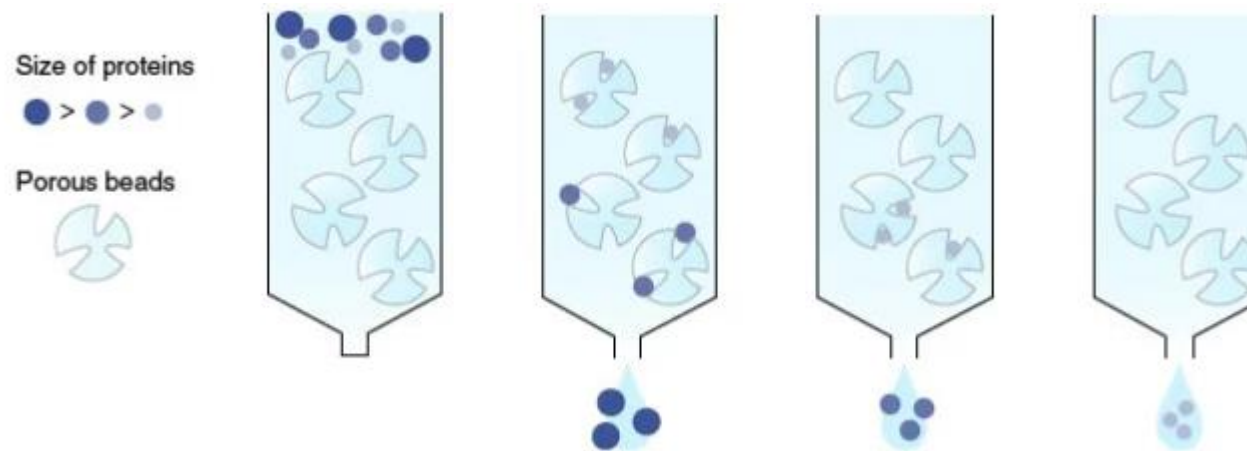


## 2.1. Separación o fraccionamiento de proteínas

### Cromatografía:

Tipos:

- **Cromatografía de exclusión molecular** – se aplica en condiciones nativas manteniendo la estructura de las proteínas. Las moléculas se separan en función de su **tamaño**. La fase estacionaria es una matriz de polímero poroso que permite la entrada de unas partículas y no de otras en función del tamaño. La elución se produce usando una solución de la misma polaridad.





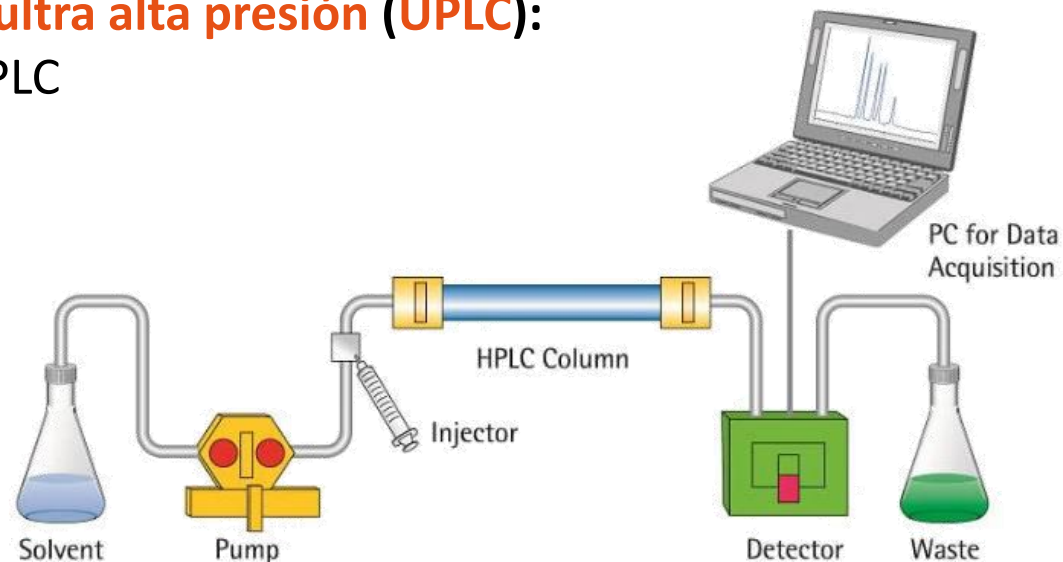
## 2.1. Separación o fraccionamiento de proteínas

### Cromatografía líquida de alta presión (HPLC):

- Recibe el nombre de alta presión porque el disolvente atraviesa la columna bajo altas presiones en lugar de por gravedad.
- Las diferentes moléculas de una mezcla tienen un grado de interacción variable con el absorbente presente en la fase estacionaria.
- Dependiendo de la afinidad de cada molécula por la fase estacionaria, su velocidad de movimiento a través de la columna varía (<afinidad >movimiento).

### \*Cromatografía líquida de ultra alta presión (UPLC):

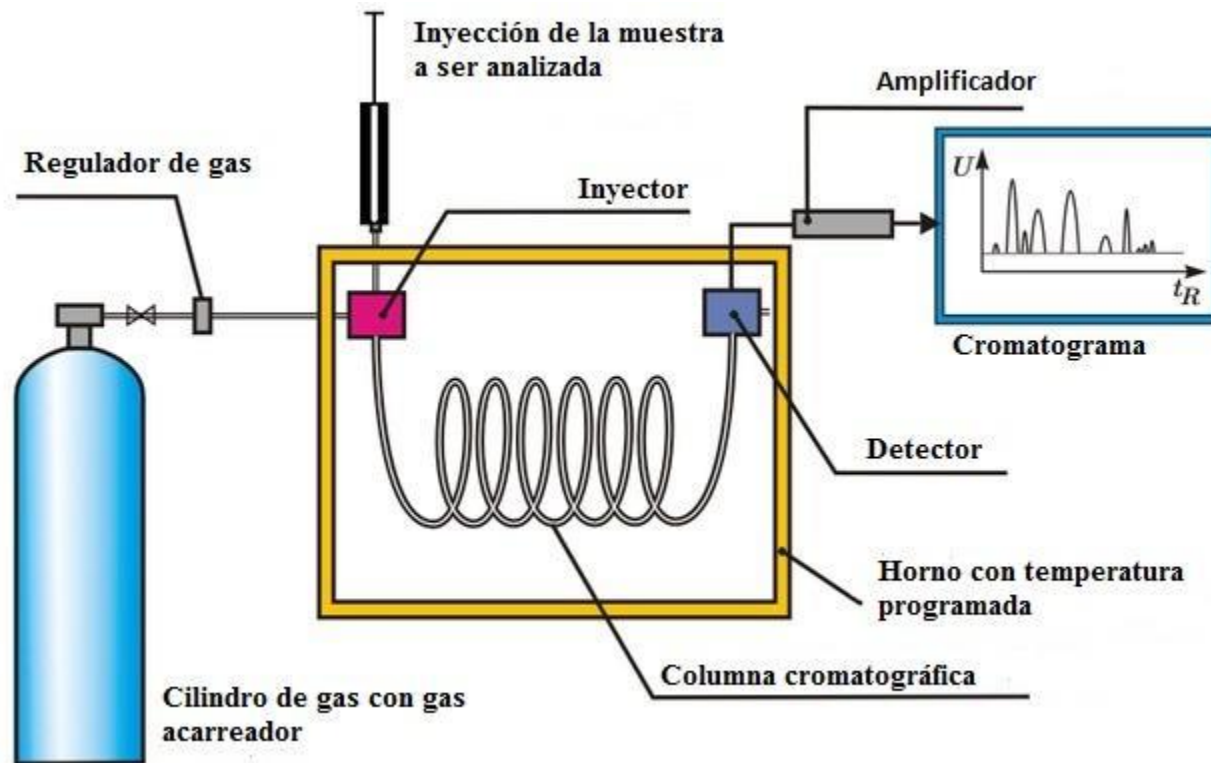
10 veces más rápida que HPLC



## 2.1. Separación o fraccionamiento de proteínas

### Clasificación según soporte de la cromatografía:

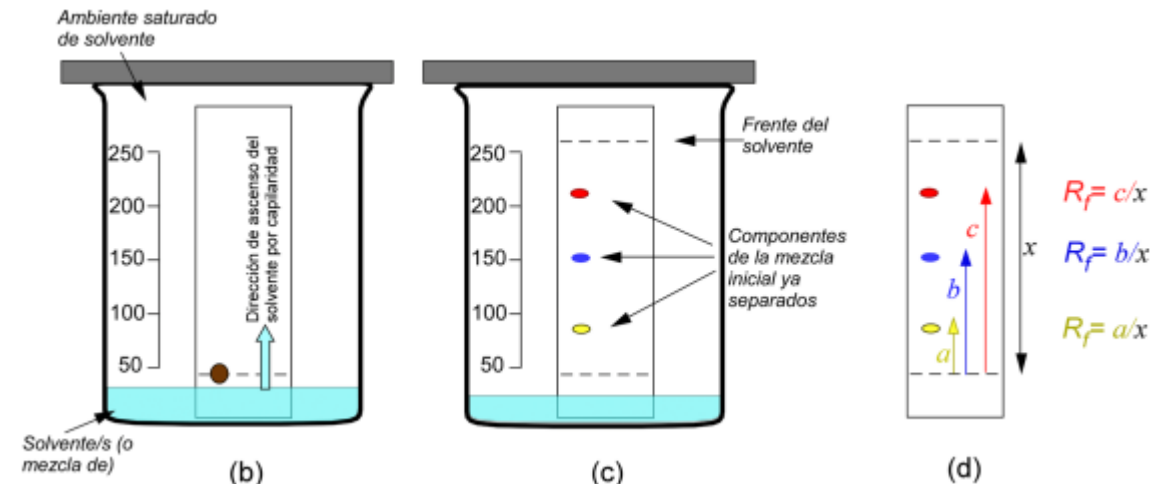
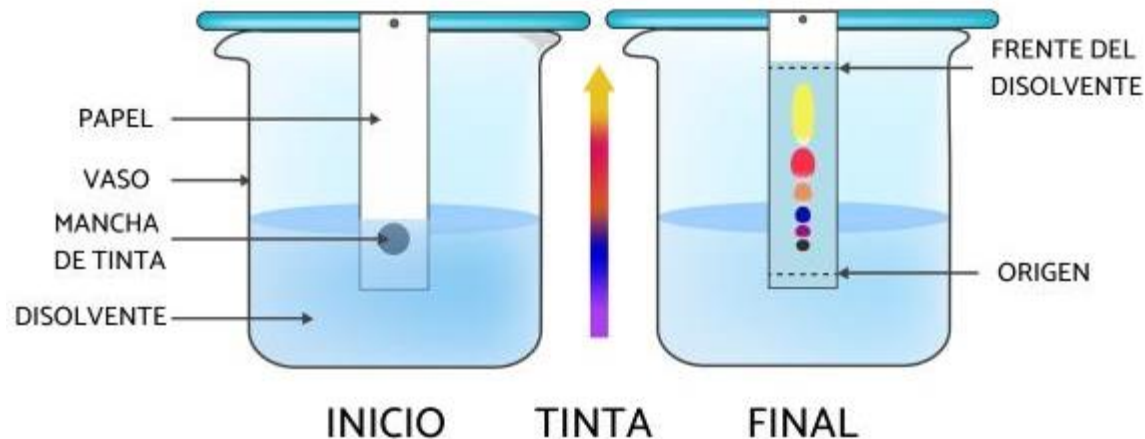
- **Cromatografía líquida (LC):** todas las anteriormente vistas
- **Cromatografía de gases (GC):** la fase estacionaria es una columna mientras que la móvil es un gas que arrastra los compuestos volátiles.



## 2.1. Separación o fraccionamiento de proteínas

### Clasificación según soporte de la cromatografía:

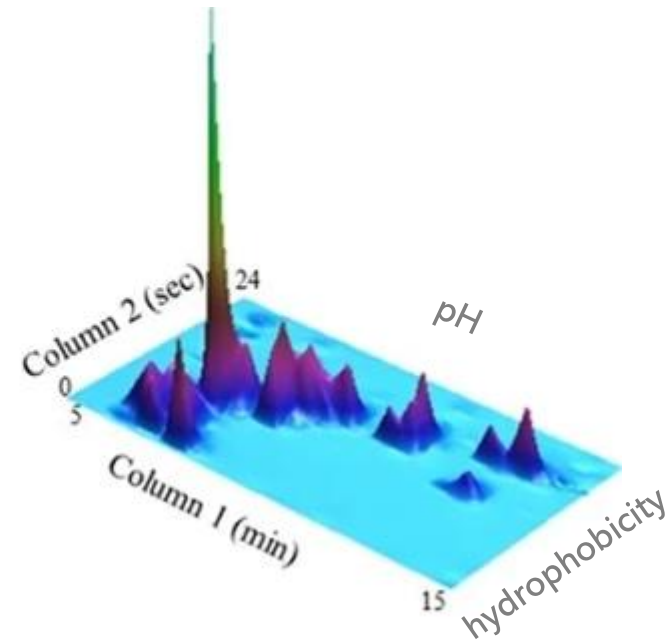
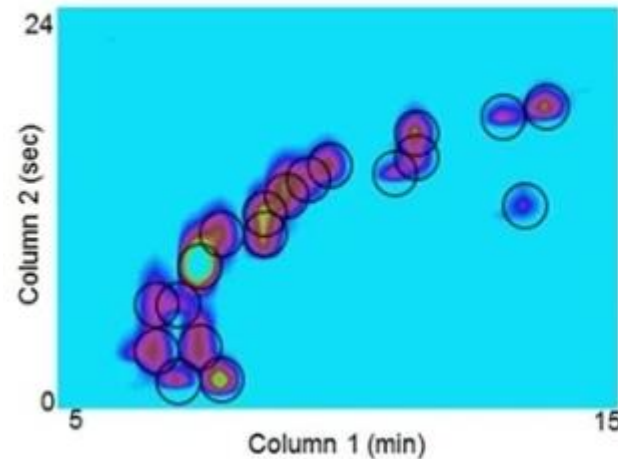
- **Cromatografía en papel:** fase estacionaria – papel filtro; fase móvil – solvente que asciende por capilaridad
- **Cromatografía en capa fina (TLC):** fase estacionaria – capa fina de silicio; fase móvil – solvente que asciende por capilaridad. Los componentes de la muestra avanzan según polaridad.



## 2.1. Separación o fraccionamiento de proteínas

### Cromatografía líquida **multidimensional**:

- Hace uso de dos o más columnas **acopladas** para conseguir separaciones no alcanzables con un único sistema.
- LC-LC, GC-GC, LC-GC.



# Temario - Contenidos

Tema 2. Métodos en proteómica: caracterización de la expresión de proteínas

2.1. Separación o fraccionamiento de proteínas

**2.2. Caracterización e identificación de proteínas**

2.3. Cuantificación y expresión diferencial de proteínas

2.4. Análisis computacional de datos en proteómica

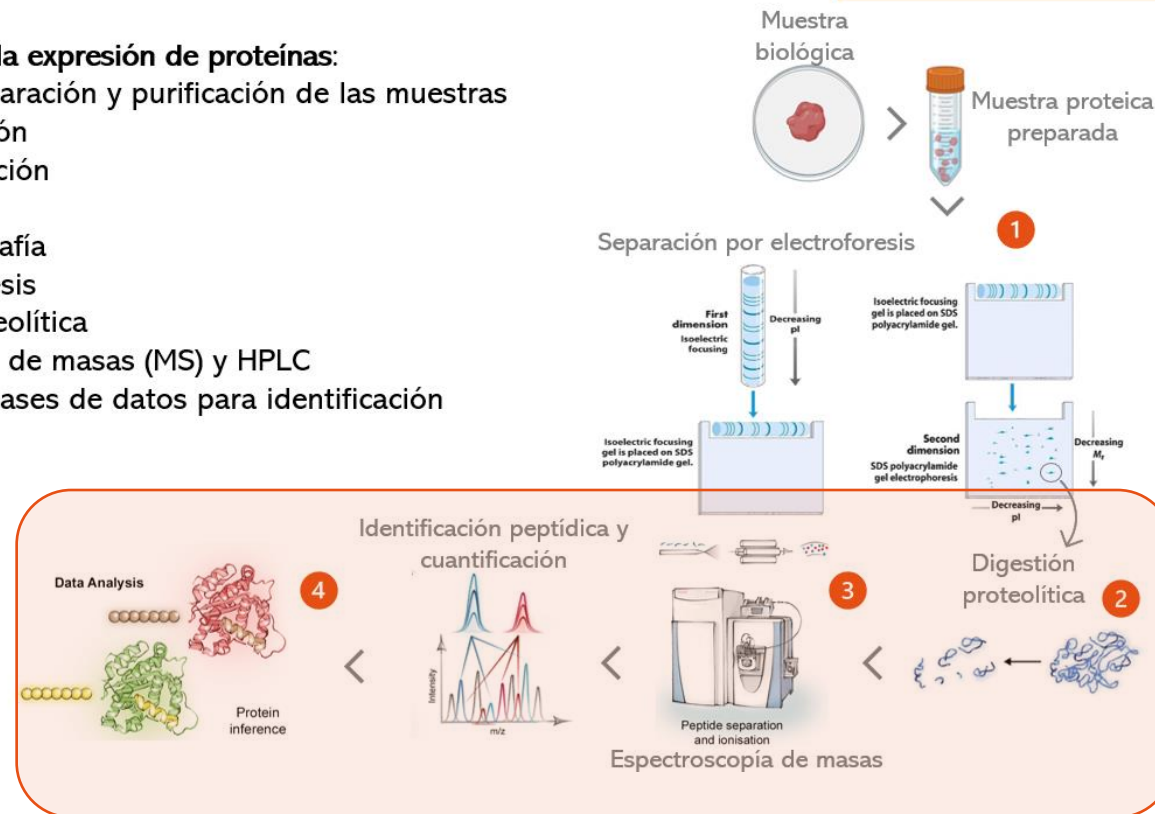
# 2.2. Caracterización e identificación de proteínas

## Espectrometría de masas:

### 1.4. Proteómica - metodologías

Caracterización de la expresión de proteínas:

1. Extracción, separación y purificación de las muestras
  - Precipitación
  - Centrifugación
  - Filtración
  - Cromatografía
  - Electroforesis
2. Digestión proteolítica
3. Espectrometría de masas (MS) y HPLC
4. Búsqueda en bases de datos para identificación



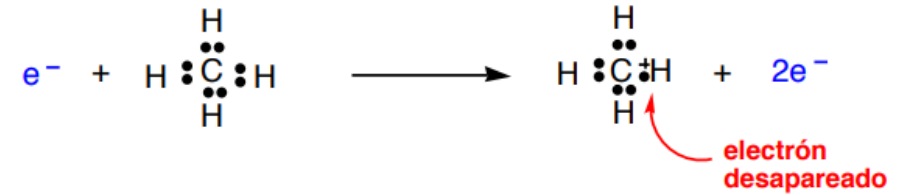
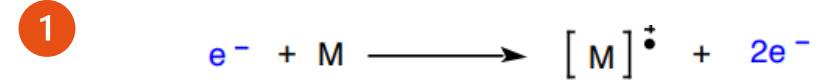
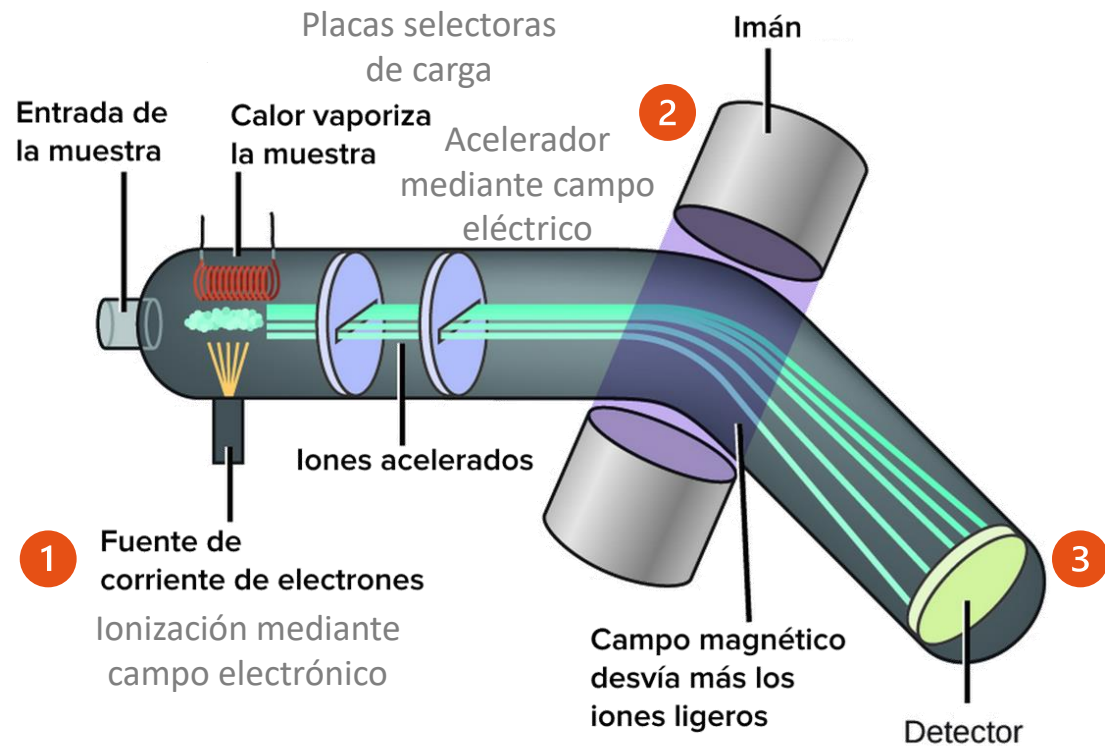
## 2.2. Caracterización e identificación de proteínas

### Espectrometría de masas:

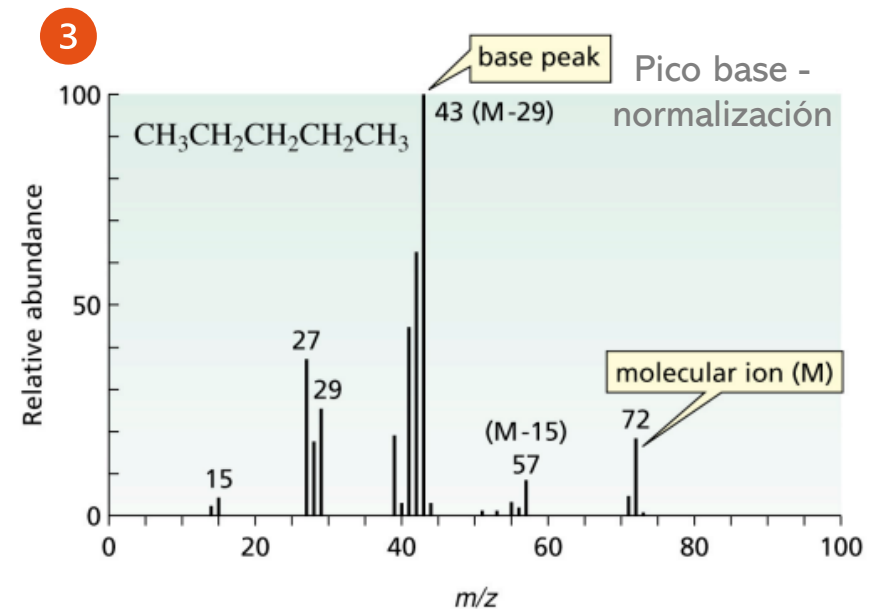
- Se basa en la separación en vacío de iones en fase gaseosa de acuerdo con su relación masa/carga ( $m/z$ ).
- No confundir **espectroscopía (ES)** con **espectrometría (EM)**.
- **Espectroscopía**: medimos la radiación que absorbe o emite la muestra (fotones).
- **Espectrometría**: se bombardea con electrones proporcionando una energía superior a la necesaria para romper enlaces. Se produce una fragmentación de la molécula por los enlaces más débiles posibilitando su identificación.

## 2.2. Caracterización e identificación de proteínas

### Espectrometría de masas:



cation-radical



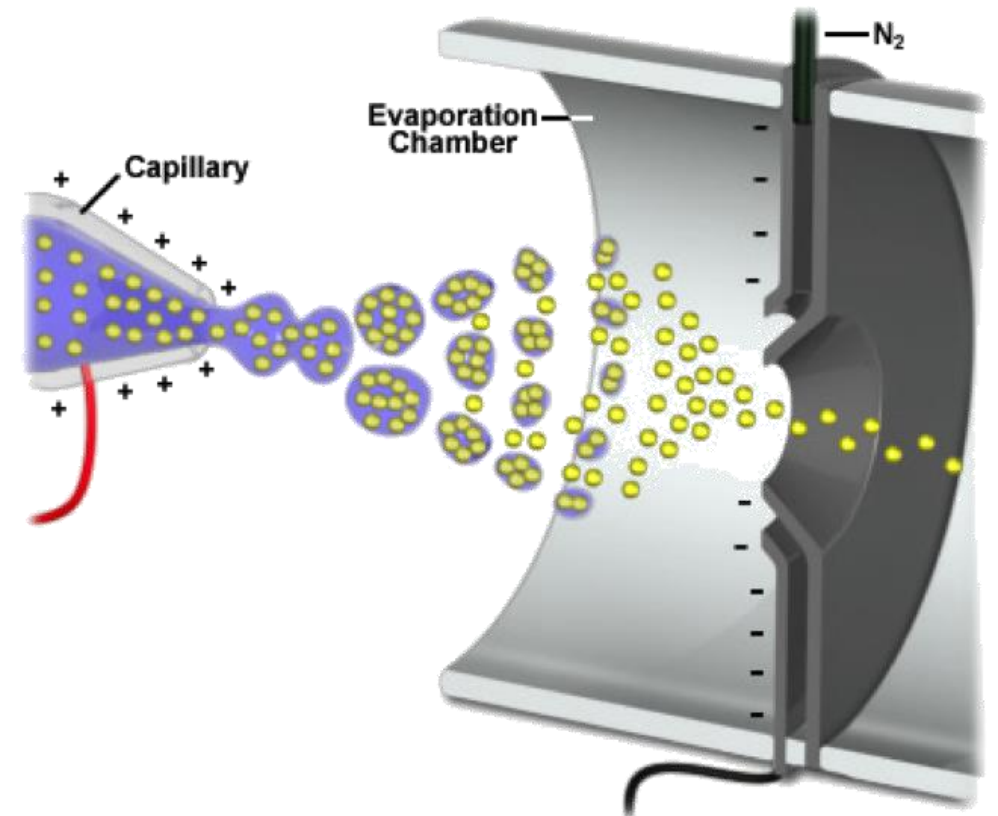


## 2.2. Caracterización e identificación de proteínas

### Espectrometría de masas:

#### 1. Sistemas de ionización:

- Fase gaseosa – primero se volatiliza la muestra y después se ioniza
  - **Impacto electrónico (EI)** – el vapor molecular se ioniza al bombardearse con un haz de  $e^-$
  - **Ionización química (CI)** – se bombardea con  $e^-$  un gas reactivo que colisiona con la muestra ionizándola
- Desorción – la muestra sólida o líquida se transforma directamente en iones
  - **Ionización por electrospray (ESI)** – la muestra disuelta en disolvente volátil pasa por un fino capilar de acero cuyo extremo está sometido a un potencial eléctrico elevado. Se producen reacciones de oxido-reducción cargando eléctricamente las gotas.

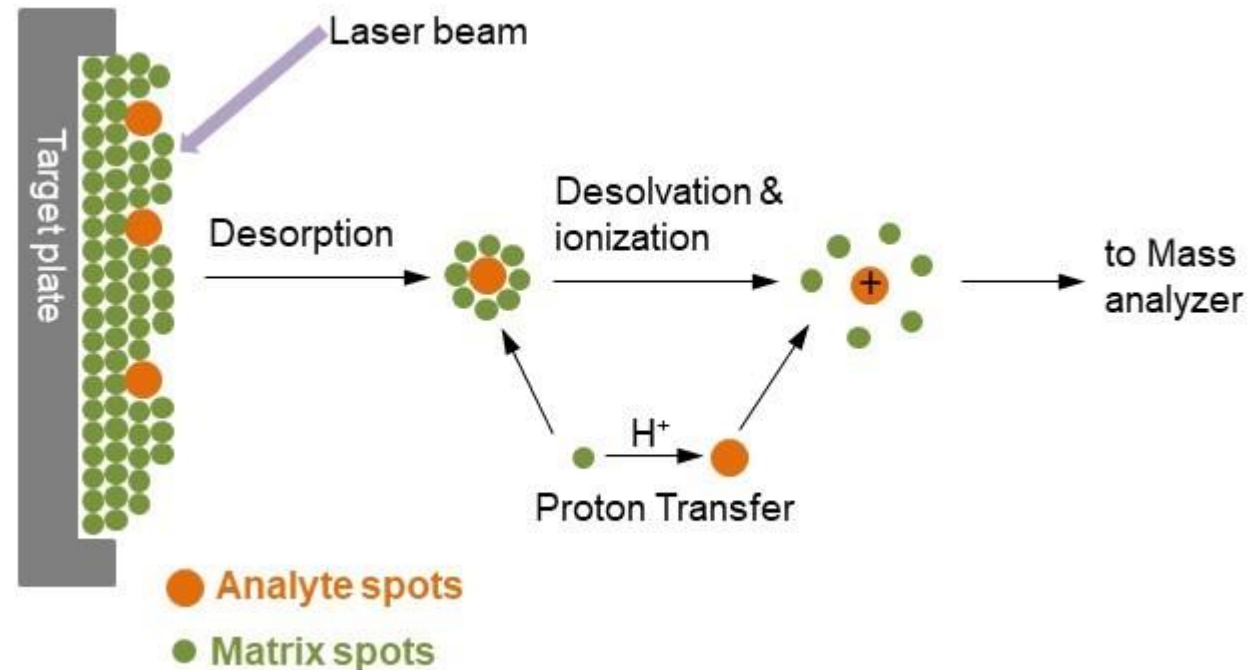


## 2.2. Caracterización e identificación de proteínas

### Espectrometría de masas:

#### 1. Sistemas de ionización:

- Desorción – la muestra sólida o líquida se transforma directamente en iones
  - **Asistida por matriz (MALDI)** – la muestra se mezcla con disolvente que contiene compuestos orgánicos pequeños (matriz) y que presentan una absorción elevada a  $\lambda$  del láser. Se deja cristalizar. Se aplican pulsos del láser sobre el cristal produciendo su **sublimación** y transferencia de  $H^+$  a las moléculas de la muestra.

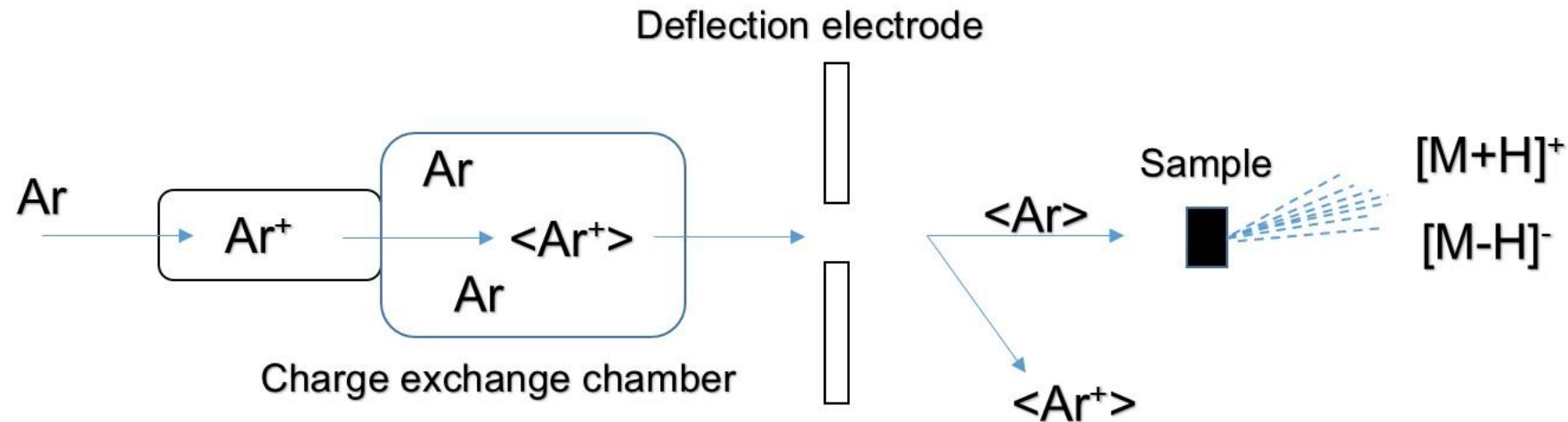


## 2.2. Caracterización e identificación de proteínas

### Espectrometría de masas:

#### 1. Sistemas de ionización:

- Desorción – la muestra sólida o líquida se transforma directamente en iones
  - **Bombardeo con átomos rápidos (FAB)** – la muestra se ioniza mediante bombardeo con átomos de Xe o Ar de elevada energía.



## 2.2. Caracterización e identificación de proteínas

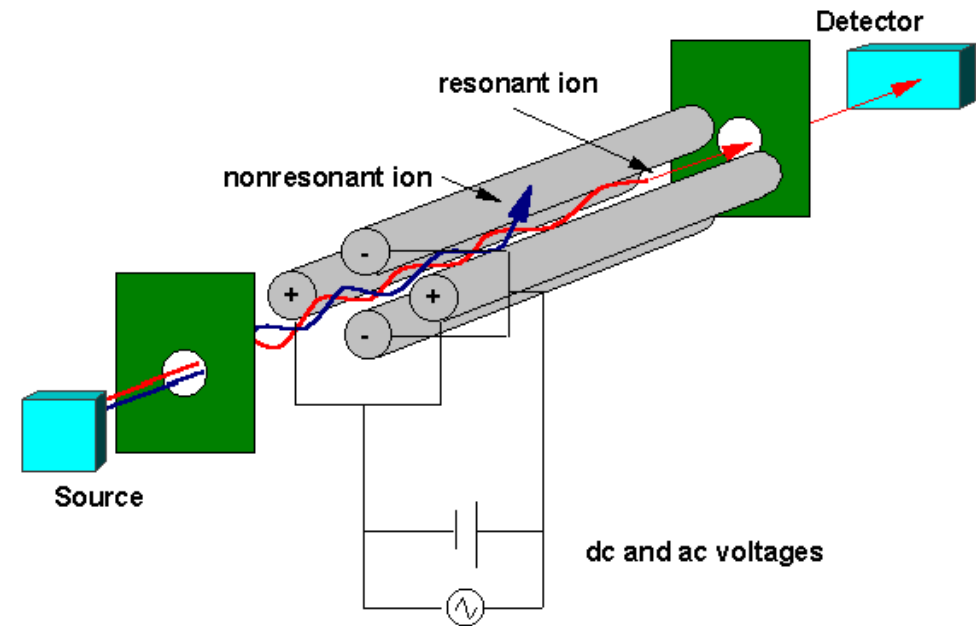
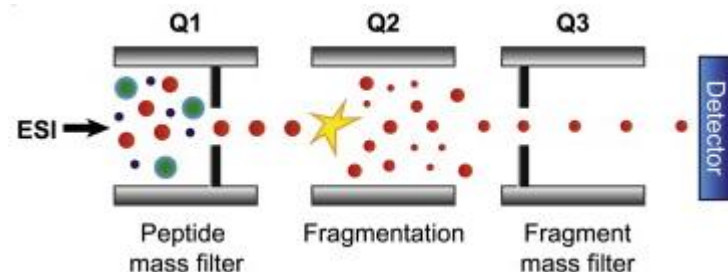
### Espectrometría de masas:

#### 2. Analizadores de masas:

- **Sector magnético** – un campo magnético afecta al radio de curvatura de los iones
- **Sector eléctrico** – un campo eléctrico afecta al radio de curvatura de los iones
- **Cuadrupolo (Q)** – barrido en un campo de radiofrecuencias.

A las varillas opuestas se les aplica un potencial electrostático continuo de igual signo y opuesto al de las otras dos varillas. Funciona como un filtro de masas.

*Full Scan* (barrido) o *Selected Ion Monitoring*  
Triplecuadrupolo (QqQ)

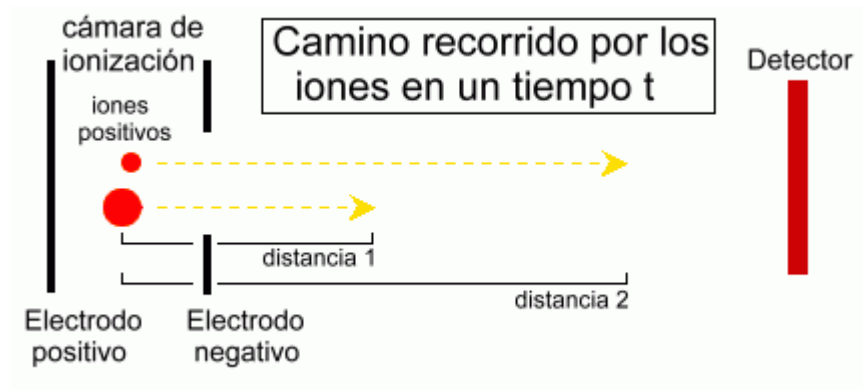
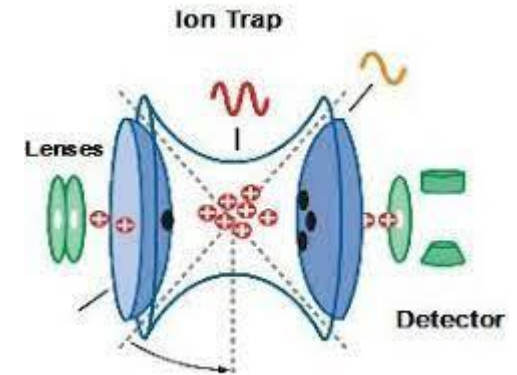


## 2.2. Caracterización e identificación de proteínas

### Espectrometría de masas:

#### 2. Analizadores de masas:

- **Trampa de iones (IT)** – funciona igual que Q solo que tridimensional
- **Tiempo de vuelo (TOF)** – la separación de los iones según relación  $m/z$  se produce durante su recorrido hasta el detector situado al final del tubo de vuelo ya que su velocidad depende de su masa.



## 2.2. Caracterización e identificación de proteínas

### Combinación cromatografía (LC) y espectrometría de masas (MS):

Permite la separación de los componentes de la muestra y la obtención del espectro de masas correspondiente a cada componente, que puede ser empleado para elucidar su estructura o bien identificar la proteína en cuestión.

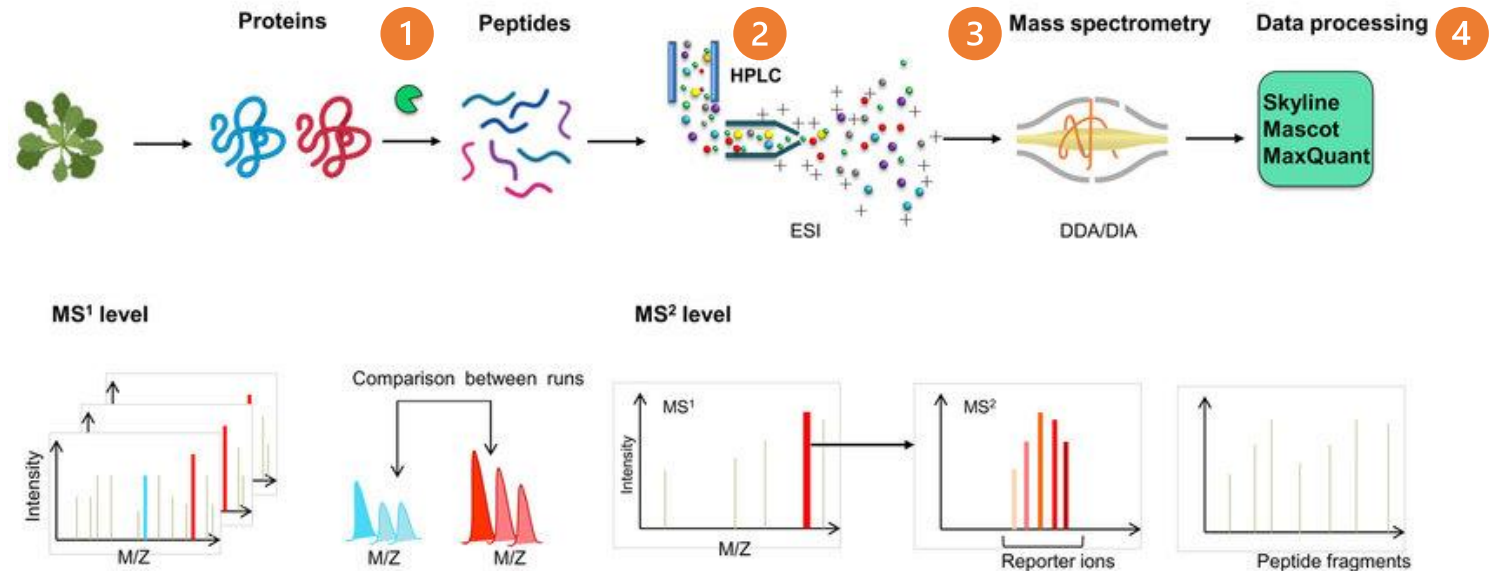
LC permite enviar a MS péptidos suficientemente separados a una alta concentración.

### Proteómica *shotgun*:

Análisis directo de una mezcla proteica compleja para generar rápidamente un perfil global.

#### Ventajas:

- Identificación rango picomolar
- Bajas pérdidas de muestra (gel)
- Digestión proteolítica proporciona péptidos identificables



# Temario - Contenidos

Tema 2. Métodos en proteómica: caracterización de la expresión de proteínas

2.1. Separación o fraccionamiento de proteínas

2.2. Caracterización e identificación de proteínas

**2.3. Cuantificación y expresión diferencial de proteínas**

2.4. Análisis computacional de datos en proteómica

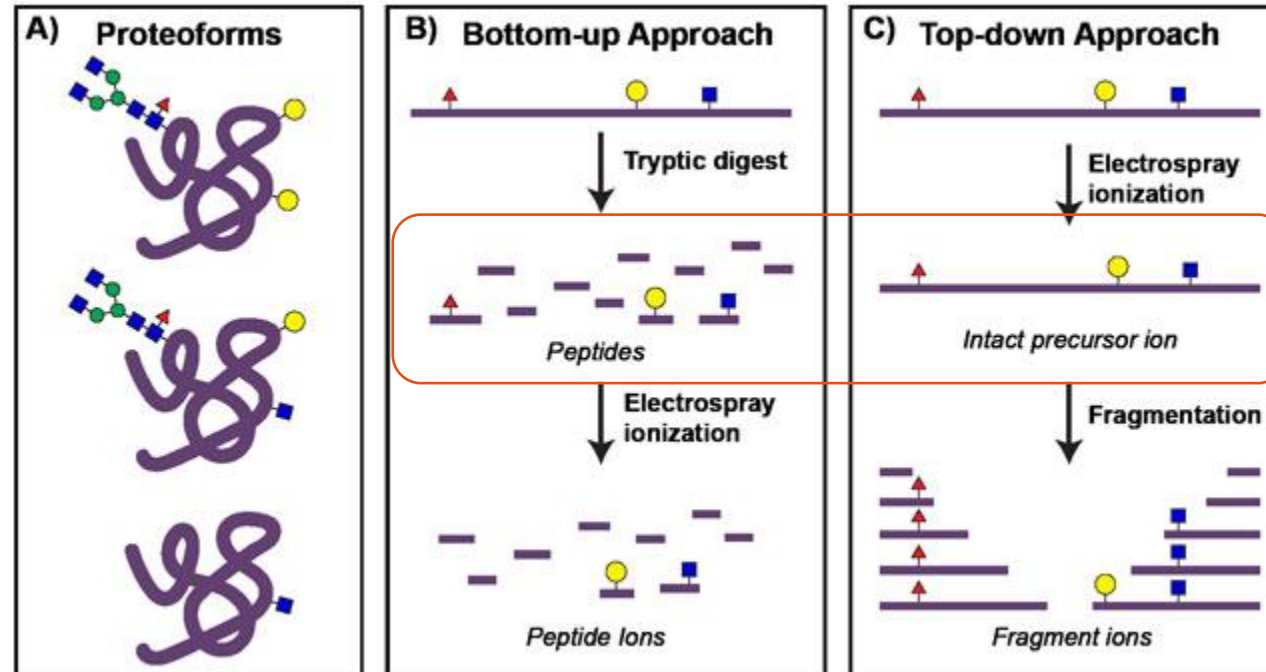


## 2.3. Cuantificación y expresión diferencial de proteínas

### Proteómica *Top Down & Bottom Up*

Ventajas BU:

- Mejor adaptación instrumental y de software.
- Mayor sensibilidad que TD.
- Protocolos preestablecidos y replicables.



Ventajas TD:

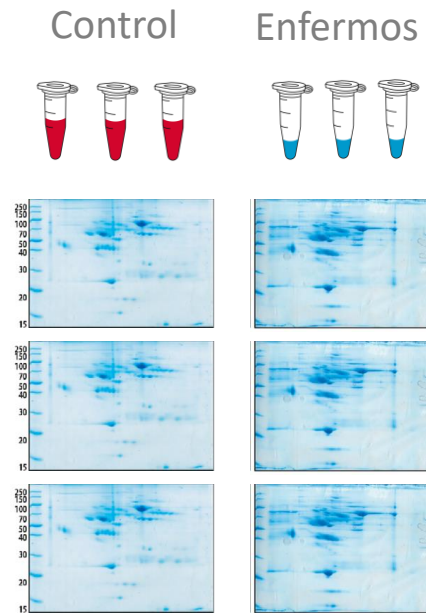
- La masa exacta de cada proteoforma es conservada.
- Se conservan características estructurales lábiles.
- Identifica mejor proteínas de bajo PM.
- Se puede emplear para detectar isoformas, mutantes, productos de degradación de N y C-terminales, modificaciones post traduccionales, productos de degradación o combinaciones de todas las anteriores.
- Proporciona estequiometría.



## 2.3. Cuantificación y expresión diferencial de proteínas

### Expresión diferencial de proteínas

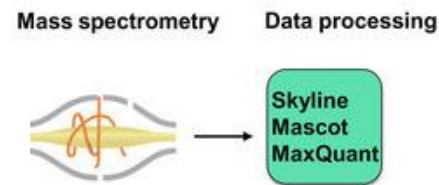
#### Aproximación clásica:



Comparación  
Cuantificación  
Selección banda

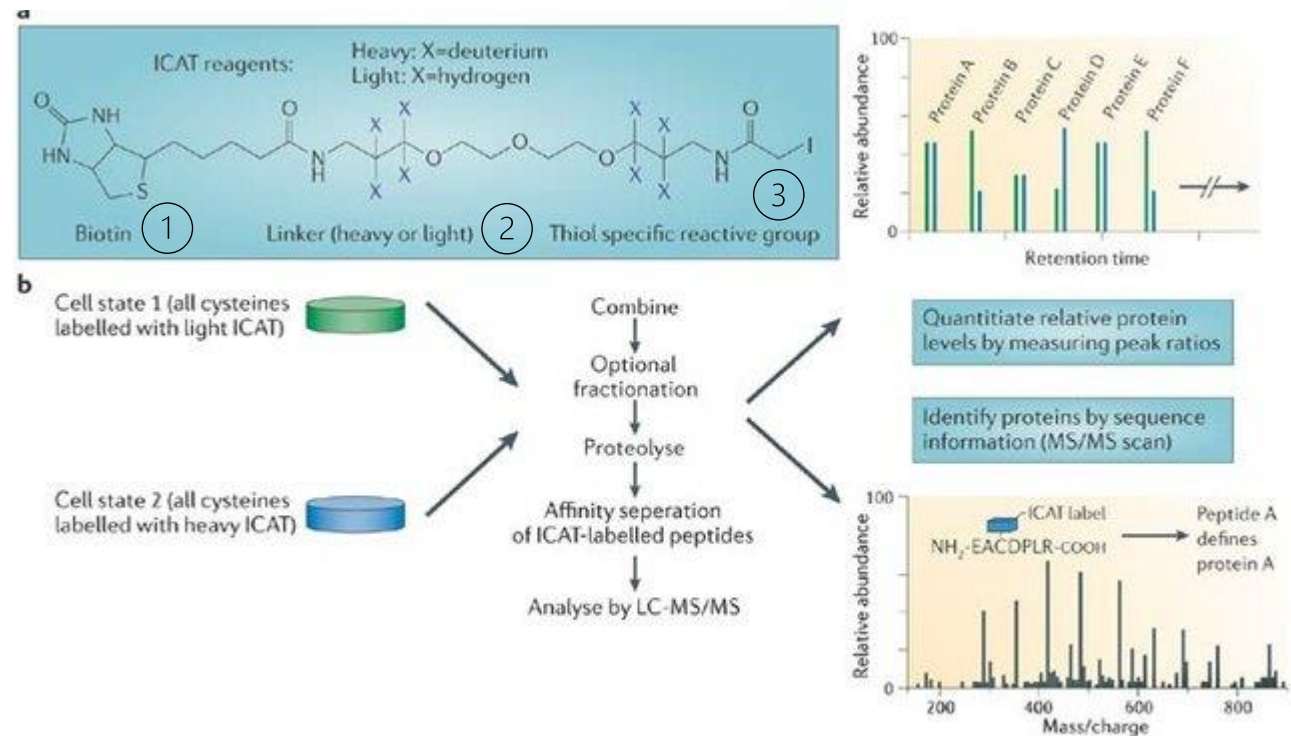


Digestión de la  
banda



#### Aproximación moderna:

### ICAT – Isotope-Coded Affinity Tag



Copyright © 2006 Nature Publishing Group  
Nature Reviews | Microbiology

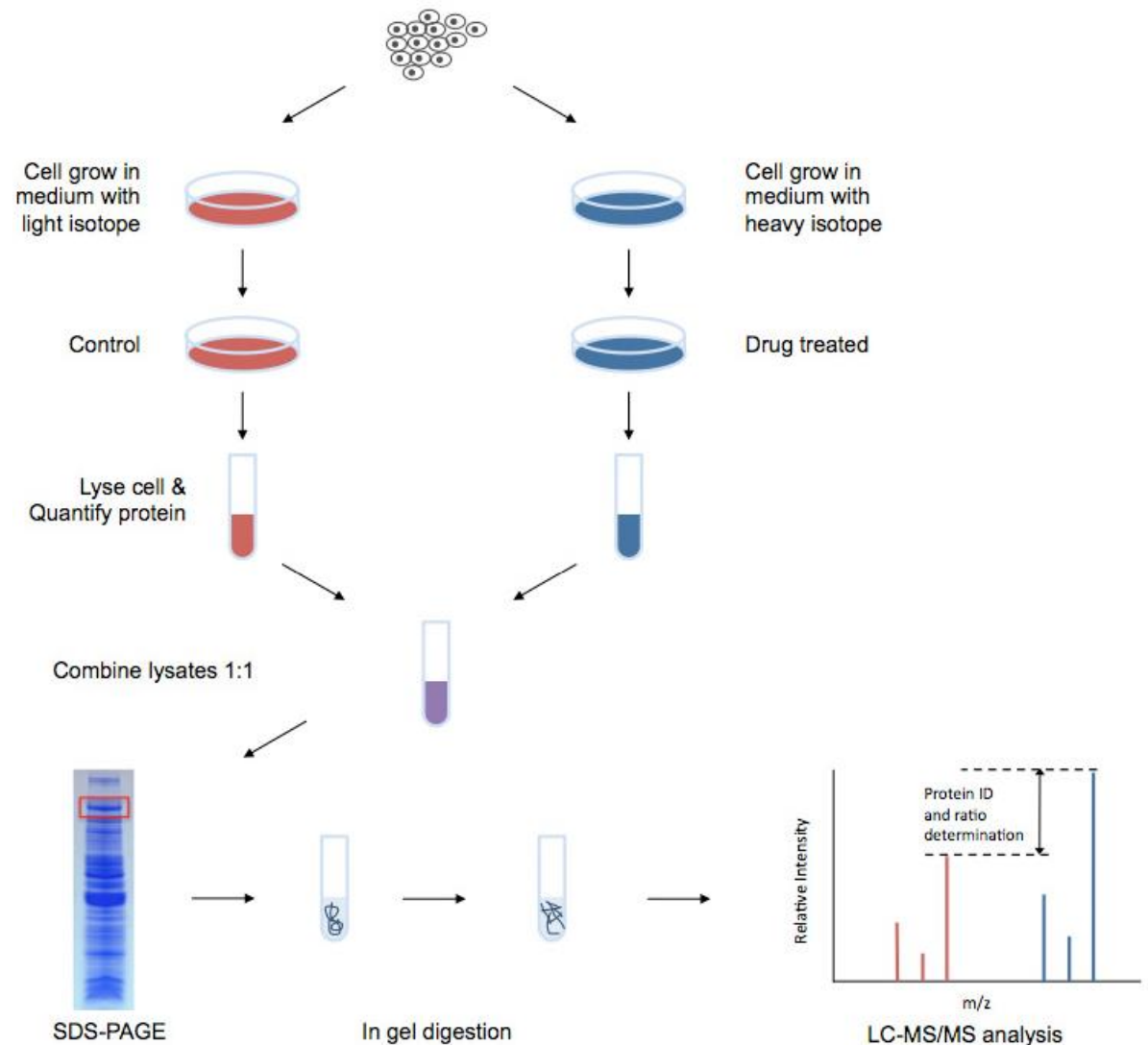
doi: 10.1038/nrmicro1552

## 2.3. Cuantificación y expresión diferencial de proteínas

### Expresión diferencial de proteínas

Aproximación moderna:

**SILAC** – Stable Isotope Labeling  
by Amino Acids in Cell Culture



## 2.3. Cuantificación y expresión diferencial de proteínas

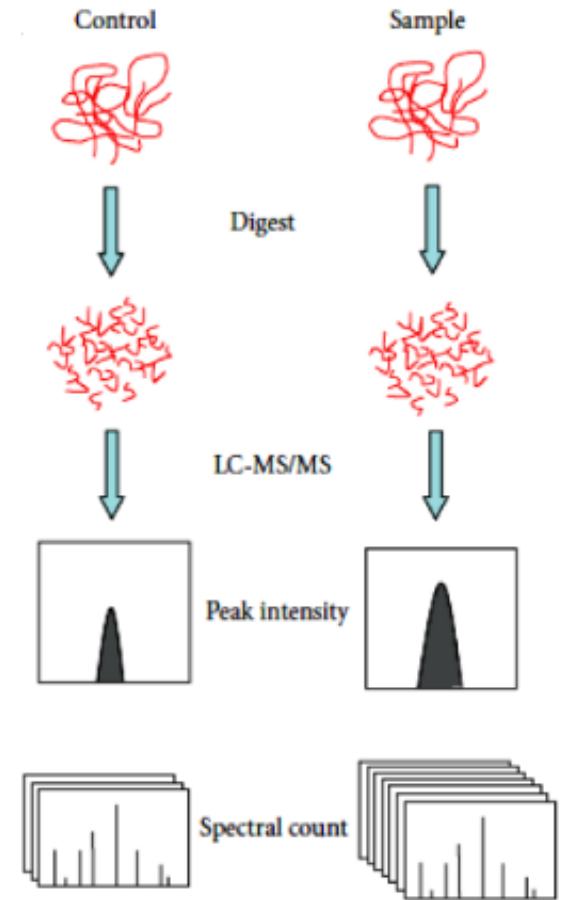
### Expresión diferencial de proteínas

Sin etiquetas (*label-free*):

¿Cómo tiene lugar la cuantificación proteica?

Existen dos formas:

- Comparando el espectro de masas – cambio en la **intensidad de señal** (área de pico) que se correlaciona con la cantidad de proteína.
- Contaje de espectros – el **número de péptidos contabilizados** para una proteína dada es indicativo de la cantidad de dicha proteína.



doi: 10.3390/proteomes1030180

# Temario - Contenidos

Tema 2. Métodos en proteómica: caracterización de la expresión de proteínas

2.1. Separación o fraccionamiento de proteínas

2.2. Caracterización e identificación de proteínas

2.3. Cuantificación y expresión diferencial de proteínas

2.4. Análisis computacional de datos en proteómica

## 2.4. Análisis computacional de datos en proteómica

### Bases de datos de proteínas y proteómica

UniProtKB – [www.uniprot.org](http://www.uniprot.org)

UNIMOD – [www.unimod.org](http://www.unimod.org); doi: 10.1002/pmic.200300744

Base de datos de **modificaciones proteicas** para su uso en aplicaciones de espectrometría de masas, especialmente para identificación proteica y secuenciación *de novo*.

UNIMOD

protein modifications for mass spectrometry

Help

Unimod, View record [ Accession #: 512 ]

Back to list

Accession #	512	PSI-MS Name	Hex(2)	Interim Name	Gal-Glu
Description	Lactosylation				
Alt. Description	lac				
Composition	Hex(2)	Monoisotopic	324.105647	Average	324.2812
Specificity Definition 1					
Site	K	Position	Anywhere	Classification	Other glycosylation
Comment	Maillard reaction (first event)				
Specificity Definition 2					
Site	R	Position	Anywhere	Classification	Other glycosylation
Comment	Maillard reaction (first event)				
Specificity Definition 3					
Site	S	Position	Anywhere	Classification	O-linked glycosylation
Neutral Loss	Hex(2)	Monoisotopic	324.105647	Average	324.2812
Neutral Loss		Monoisotopic	0.000000	Average	0.0000
Specificity Definition 4					
Site	T	Position	Anywhere	Classification	O-linked glycosylation
Neutral Loss	Hex(2)	Monoisotopic	324.105647	Average	324.2812
Neutral Loss		Monoisotopic	0.000000	Average	0.0000
Notes and References					
Source	Journal	Reference	BIOCHIMECAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS 236, 413-417 (1997)		
Source	Journal	Reference	ANALYTICAL BIOCHEMISTRY 259, 152-161 (1998)		
Source	Misc. URL	Reference	Composition search GlyConnect		
Notes	Lactosylation of bovine Beta-Lactoglobulin				
Curator	molle	Last Modified	2017-11-23 12:59:14	Verified	No

# 2.4. Análisis computacional de datos en proteómica

## Bases de datos de proteínas y proteómica

PeptideAtlas – [www.peptideatlas.org](http://www.peptideatlas.org)

# PeptideAtlas



Log In

Search PeptideAtlas:

[Expanded Search](#)

PeptideAtlas is a multi-organism, publicly accessible compendium of peptides identified in a large set of tandem mass spectrometry proteomics experiments. Mass spectrometer output files are collected for human, mouse, yeast, and several other organisms, and searched using the latest search engines and protein sequences. All results of sequence and spectral library searching are subsequently processed through the [Trans Proteomic Pipeline](#) to derive a probability of correct identification for all results in a uniform manner to insure a high quality database, along with false discovery rates at the whole atlas level. Results may be queried and browsed at the PeptideAtlas web site. The raw data, search results, and full builds can also be downloaded for other uses.

### Human Builds



Human all  
2023-01



Human Plasma  
2023-04

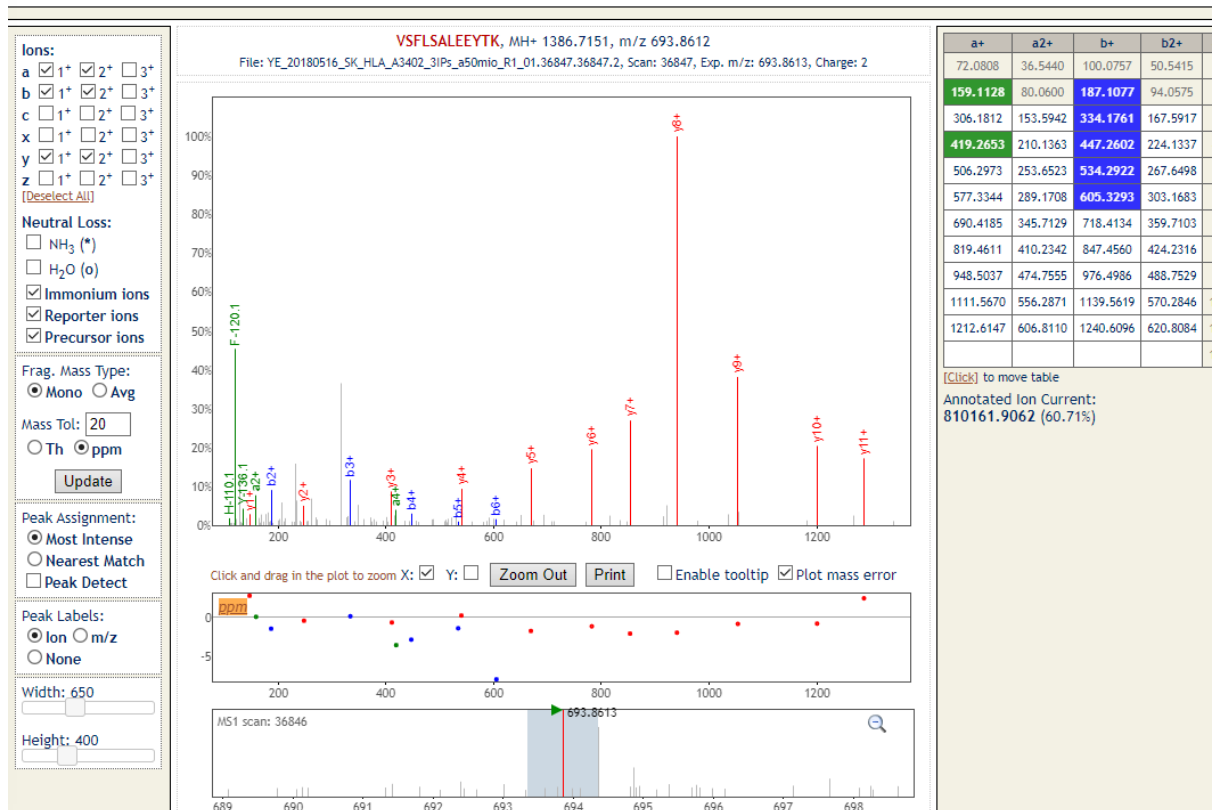


Human Plasma  
Extracellular Vesicle  
2023-04



Human Urine  
2022-02

### Spectrum for VSFLSALEEYTK<sup>+2</sup>



## 2.4. Análisis computacional de datos en proteómica

### Bases de datos de proteínas y proteómica

ProteinProspector (<https://prospector.ucsf.edu/prospector/mshome.htm>)



### ProteinProspector

v 6.4.9

*Proteomics tools for mining sequence databases  
in conjunction with Mass Spectrometry experiments.*

#### ProteinProspector Tools

#### Administration/Help

##### Batch MSMS Database Searching [Instructions](#)

[Search Compare](#) [Batch-Tag Web](#) [Batch-Tag](#)  
[Results Management](#) [Search Table](#)

##### Database Search Programs

[MS-Fit](#) [MS-Tag](#) [MS-Homology](#) [MS-Bridge](#)  
[MS-Fit Upload](#) [MS-Seq](#) [MS-Pattern](#) [MS-NonSpecific](#)

##### Peptide / Protein MS Utility Programs

[MS-Digest](#) [MS-Product](#) [MS-Filter](#) [MS-Viewer](#)  
[MS-Isotope](#) [MS-Comp](#)

##### Database Management

[DB-Stat](#)

##### [Video Tutorials](#)

[Administering ProteinProspector](#)

[User's Manual](#)

[FAQ](#)

[Bug Listing](#)

[ProteinProspector Revision History](#)

[ProteinProspector Automation Guidance](#)

[ProteinProspector Local Installation](#)

##### Useful Tables

- [Mutation Mass Shifts](#)
- [Dipeptide Masses](#)
- [Trypsin Autolysis Products](#)
- [Glycosylation Modification Masses](#)
- [Averagine Distributions](#)

##### [Publications](#)

[MS-Viewer Repository Browser](#)

[Useful Links](#)



# 2.4. Análisis computacional de datos en proteómica

## Bases de datos de proteínas y proteómica

ProteinProspector ([www.prospector.ucsf.edu/prospector/mshome.htm](http://www.prospector.ucsf.edu/prospector/mshome.htm))

## Simulador de digestión proteolítica

### MS-Digest

Database: UniProtKB.2020.09.02  
Output: HTML Hits to file ☐ Name: lastres  
Digest: Trypsin Max. Missed Cleavages: 1  
[+] End Terminus Parameters  
Abu (N-term)  
Acetohydrazide (C-term)  
Acetohydrazide (DE)  
Constant Mods  
Acetyl (K)  
Acetyl (N-term)  
Acetyl (R)  
[+] Present Amino Acids (see instructions)  
Perform Digest  
Variable Mods  
Acetyl (Protein N-term)  
Gln->pyro-Glu (N-term Q)  
Met-loss (Protein N-term M)  
Met-loss+Acetyl (Protein N-term M)  
Oxidation (M)  
Add (+) or remove (-) mods using menus & buttons below  
Frequent  
Acetyl (K)  
Acetyl (Protein N-term)  
Acetyl (Uncleaved K)  
Acetyl+Oxidation (Protein N-term M)  
[+] Set Contents  
For digestion of a user supplied sequence select User Protein above.  
User Protein Sequence  
MKAVVLTAVLFLTGSQARHFQQDDPQSSWDRVKDFATVYVEAIKDSGRDYVAQFEASALGKQLNLKLLDNWDTLASTL  
SKVREQLGPVTQEFWDNLEKETASLRQEMHKDL EEVKQVQPYLDEFQKKWHEEVEIYRQKVAPLGEFFREGARQKVQEL  
QDKLSPLAQELRDRARAHVETLRQQLAPYSDDL RQRLTARLEALKEGGSLAEYHAKASEQLKALGEKAKPVLEDLRQGL  
LPVLES LKVSILAAIDEASKKLNQ

Index Number: 152195

pI of Protein: 7.0

Protein MW: 15733

Amino Acid Composition: A21 C1 D9 E4 F9 G8 H3 I8 K11 L12 M4 N7 P1 Q6 R2 S9 T12 V10 W3 Y4

1 VSANDIKNVQ DTWGKLYDQW DAVHASKFYN KLFKDESDIS EAFVKAGTGS GIAMKQALV FGAILQEFVA NLNDPTALT  
81 KIKGLCATHK TRGITNMELF AFALADLVAY MGTTISFTAA QASWTAVND VILHQMSSYF ATVA

Number	m/z (mi)	m/z (av)	Modifications	Start	End	Missed Cleavages	Sequence
1	892.4557	893.0559		46	55	0	(K)AGTGSGIAMK(R)
1	908.4506	909.0553	1Oxidation	46	55	0	(K)AGTGSGIAMK(R)
1	947.4581	948.0292		8	15	0	(K)NVQDTWGK(L)
1	959.5349	960.1710		28	34	1	(K)FYNKLFK(D)
1	970.5502	971.2170		82	90	1	(K)IKGLCATHK(T)
1	986.5200	987.1757		84	92	1	(K)GLCATHKTR(G)
1	1048.5568	1049.2445		46	56	1	(K)AGTGSGIAMKR(Q)
1	1064.5517	1065.2439	1Oxidation	46	56	1	(K)AGTGSGIAMKR(Q)
1	1239.5739	1240.3170		35	45	0	(K)DSEDISFAFVK(A)
1	1432.6856	1433.5729		16	27	0	(K)LYDQWDAVHASK(F)
1	1627.8214	1628.8305		32	45	1	(K)LFKDESDISFAFVK(A)
1	1674.8446	1675.8499		1	15	1	(-)VSANDIKNVQDTWGK(L)
1	1716.8551	1717.8874	1Acetyl	1	15	1	(-)VSANDIKNVQDTWGK(L)
1	1984.9552	1986.2076		16	31	1	(K)LYDQWDAVHASKFYNK(L)
1	2113.0118	2114.3502		35	55	1	(K)DSEDISFAFVKAGTGS IAMK(R)
1	2129.0067	2130.3495	1Oxidation	35	55	1	(K)DSEDISFAFVKAGTGS IAMK(R)
1	2361.1258	2362.5794		8	27	1	(K)NVQDTWGKLYDQWDAVHASK(F)
1	2411.1700	2412.7434		123	144	0	(K)ASWTAVNDVILHQMSSYFATVA(-)
1	2427.1649	2428.7428	1Oxidation	123	144	0	(K)ASWTAVNDVILHQMSSYFATVA(-)
1	2669.4549	2671.1284	1Gln->pyro-Glu	57	81	0	(R)QALVFGAILQEFVANLNDPTALT(L)
1	2686.4814	2688.1590		57	81	0	(R)QALVFGAILQEFVANLNDPTALT(L)
1	2842.5825	2844.3477		56	81	1	(K)RQALVFGAILQEFVANLNDPTALT(L)
1	2910.6339	2912.4640	1Gln->pyro-Glu	57	83	1	(R)QALVFGAILQEFVANLNDPTALT(L)
1	2927.6605	2929.4947		57	83	1	(R)QALVFGAILQEFVANLNDPTALT(L)
1	3195.6105	3197.7633		93	122	0	(R)GITNMELFAFALADLVAYMGTTISFTAAQK(A)
1	3211.6054	3213.7627	1Oxidation	93	122	0	(R)GITNMELFAFALADLVAYMGTTISFTAAQK(A)
1	3227.6003	3229.7620	2Oxidation	93	122	0	(R)GITNMELFAFALADLVAYMGTTISFTAAQK(A)
1	3452.7593	3455.0577		91	122	1	(K)TRGITNMELFAFALADLVAYMGTTISFTAAQK(A)
1	3468.7542	3471.0570	1Oxidation	91	122	1	(K)TRGITNMELFAFALADLVAYMGTTISFTAAQK(A)
1	3484.7491	3487.0564	2Oxidation	91	122	1	(K)TRGITNMELFAFALADLVAYMGTTISFTAAQK(A)



# 2.4. Análisis computacional de datos en proteómica

## Bases de datos de proteínas y proteómica

ProteinProspector ([www.prospector.ucsf.edu/prospector/mshome.htm](http://www.prospector.ucsf.edu/prospector/mshome.htm))

Calculador de distribución isotópica

### MS-Isotope

**Instructions For Entering Peptide Sequences**

Enter Sequence in Capital letters (B, J, O, X, Z not allowed) except:  
| m - Met-ox | h - Homoserine lactone | U - Selenocysteine |  
| s, t, y - Phosphorylated S, T, Y | d, e, f, g, h, i, j, k, l - user specified amino acids |  
Modified amino acids may be entered using PSI notation - eg. M(Oxidation), S(Phospho)  
[+] Click + to see list of available PSI modifications (enter exactly as shown)  
Modifications may be added together - eg. X(MappingN+Deamidated)  
Modified N and C termini must be selected from the menus

Charge	Intensity	Distribution Type	Value
1	1	Peptide Sequence	SAMPLER

[+] Additional Distributions

User Specified AA Elem Comp	Value
User Specified AA Elem Comp (d)	C2 H3 N1 O1
User Specified AA Elem Comp (e)	
User Specified AA Elem Comp (f)	
User Specified AA Elem Comp (g)	
User Specified AA Elem Comp (h)	
User Specified AA Elem Comp (i)	
User Specified AA Elem Comp (j)	
User Specified AA Elem Comp (k)	
User Specified AA Elem Comp (l)	

**Calculate Isotope Distribution**

Profile Type: Gaussian Resolution: 10000.0 Display Graph: ☒ Detailed Report: ☐

Output: HTML Hits to file: ☐ Name: lastres

<sup>13</sup>C%: 100 <sup>15</sup>N%: 100 <sup>18</sup>O%: 100 <sup>2</sup>H%: 100

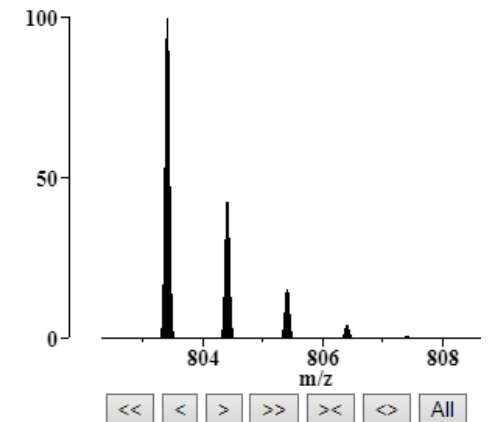
### [+] Parameters

Elemental Composition: **C33 H59 N10 O11 S1**

Monoisotopic M/Z: **803.40800**

Total Abundance: **100.00%**

Isotope Number	m/z	Percent Total	Percent Maximum
0	803.40800	61.23	100.00
1	804.41081	26.15	42.70
2	805.41062	9.54	15.58
3	806.41158	2.46	4.02
4	807.41271	0.52	0.84
5	808.41414	0.09	0.15
6	809.41572	0.01	0.02
7	810.41712	0.00	0.00
8	811.41843	0.00	0.00



# 2.4. Análisis computacional de datos en proteómica

## Bases de datos de proteínas y proteómica

Expasy – Peptide Cutter ([web.expasy.org/peptide\\_cutter](http://web.expasy.org/peptide_cutter))

## Digestión proteolítica

Expasy

PeptideCutter

Home

**PeptideCutter**

**PeptideCutter** [references / documentation] predicts potential cleavage sites cleaved by proteases or chemicals in a given protein sequence. PeptideCutter returns the query sequence with the possible cleavage sites mapped on it and/or a table of cleavage site positions.

Enter a UniProtKB (Swiss-Prot or TrEMBL) protein identifier, ID (e.g. ALBU\_HUMAN), or accession number, AC (e.g. P04406), or an amino acid sequence (e.g. 'SERVELAT'):

the cleavage of the protein.  the fields.

**Please, select**

☒ all available enzymes and chemicals

☐ only the following selection of **enzymes and chemicals**

☐ Arg-C proteinase

☐ BNPS-Skatole

☐ Caspase3

☐ Caspase6

☐ Asp-N endopeptidase

☐ Caspase1

☐ Caspase4

☐ Caspase7

☐ Asp-N endopeptidase + N-terminal Glu

☐ Caspase2

☐ Caspase5

☐ Caspase8

These enzymes cleave the sequence:

Name of enzyme	No. of cleavages	Positions of cleavage sites
Asp-N endopeptidase	1	3
Asp-N endopeptidase + N-terminal Glu	1	3
CNBr	1	5
Chymotrypsin-high specificity (C-term to [FYW], not before P)	1	12
Chymotrypsin-low specificity (C-term to [FYWML], not before P)	3	5 7 12
Formic acid	1	4
LysC	1	8
LysN	1	7
Pepsin (pH1.3)	1	12
Pepsin (pH>2)	1	12
Proteinase K	5	1 3 9 12 13
Thermolysin	2	2 11
Trypsin	1	8

- Please note that the cleavage occurs at the **right side** (C-terminal direction) of the marked amino acid.
- You have the possibility to display the results of a single enzyme by **mouseclicking** on the respective enzyme name in the map.

```
Ch_hi_Ch_lo_Pn1.3_Pn2_ProtK
  Therm|
  ProtK ||
  LysC_Tryps| ||
  Ch_lo_LysN|| ||
  CNBr_Ch_lo ||| ||
  HCOOH| ||| ||
  AspN_AspGluN_ProtK|| ||| ||
  Therm||| ||| ||
  ProtK||| ||| ||

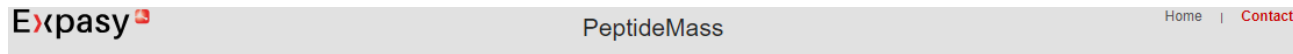
||||| ||| ||
ASIDMNHKLPGFT
1 ----- 13
```

# 2.4. Análisis computacional de datos en proteómica

## Bases de datos de proteínas y proteómica

Expasy – Peptide Mass ([web.expasy.org/peptide\\_mass](http://web.expasy.org/peptide_mass))

## Digestión proteolítica



### PeptideMass

**PeptideMass** [references] cleaves a protein sequence from the UniProt Knowledgebase (Swiss-Prot and TrEMBL) or a user-entered protein sequence with a chosen enzyme, and computes the masses of the generated peptides. The tool also returns theoretical isoelectric point and mass values for the protein of interest. If desired, PeptideMass can return the mass of peptides known to carry post-translational modifications, and can highlight peptides whose masses may be affected by database conflicts, polymorphisms or splice variants.

Instructions are available.

Enter a UniProtKB protein identifier, ID (e.g. ALBU\_HUMAN), or accession number, AC (e.g. P04406), or an amino acid sequence (e.g. 'SELVEGVIV'; you may specify post-translational modifications, but [PLEASE read this document first!](#)):

ASIDMNHKLPGFT

the fields.  the cleavage of the protein.

The peptide masses are with cysteines treated with:

- ☐ with acrylamide adducts  
☐ with methionines oxidized  
☒ [M+H]<sup>+</sup> or ☐ [M] or ☐ [M-H]<sup>-</sup> or ☐ [M+2H]<sup>2+</sup> or ☐ [M+3H]<sup>3+</sup>  
☐ average or ☒ monoisotopic.

Select an enzyme:

Allow for  missed cleavages.

Display the peptides with a mass bigger than  and smaller than  Dalton

sorted by ☒ peptide masses or in ☐ chronological order in the protein.

### PeptideMass

The entered sequence is:

<sup>10</sup>  
ASIDMNHKLP GFT

The selected enzyme is: Trypsin

Maximum number of missed cleavages (MC): 0

All cysteines in reduced form.

Methionines have not been oxidized.

Displaying peptides with a mass bigger than 500 Dalton.

Using monoisotopic masses of the occurring amino acid residues and giving peptide masses as [M+H]<sup>+</sup>.

The peptide masses from your sequence are:

[Theoretical pI: 6.79 / Mw (average mass): 1430.64 / Mw (monoisotopic mass): 1429.70]

mass	position	#MC	modifications	peptide sequence
915.4353	1-8	0		ASIDMNHK
534.2922	9-13	0		LPGFT

100.0% of sequence covered (you may modify the input parameters to display also peptides < 500 Da or > 100000000000 Da):

<sup>10</sup>  
ASIDMNHKLP GFT

in raw text format to be exported into an external application



viu

**Universidad**  
Internacional  
de Valencia

[universidadviu.com](http://universidadviu.com)

De:  
 Planeta Formación y Universidades