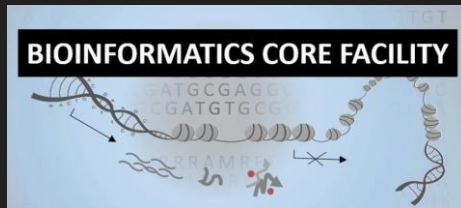


# Welcome to ABC.14

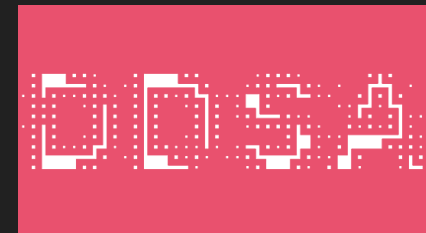
13. March 2025

<https://abc.au.dk>

abcafe@au.dk



Health  
Data Science  
Sandbox



Danish Data  
Science  
Academy

# Agenda

- What's new
- Interesting things from you?
- Topic presentation
- Tutorial and/or open coding

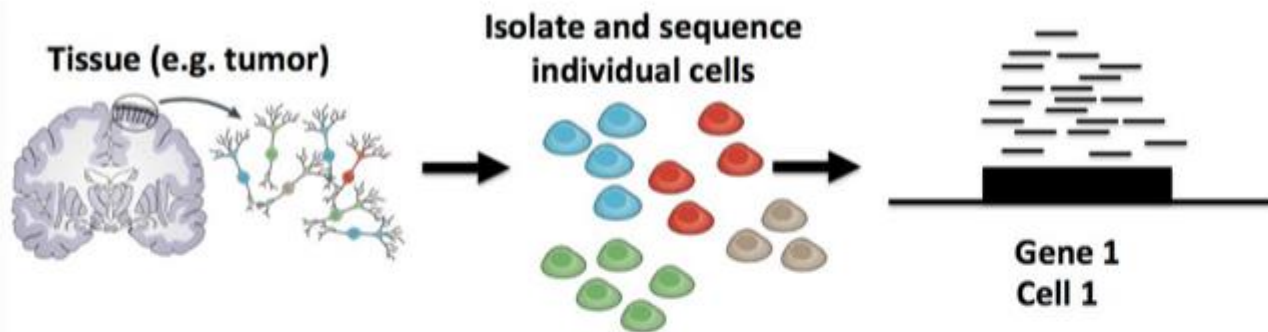
# What's new

- Still some seats available for advanced genomeDK 10th april
- Lots of seats for pipeline building (some python + bash) 4th june

# SCproteomics

## From transcription to translation

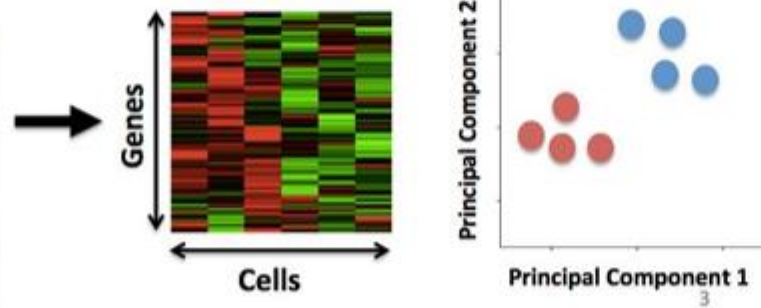
### Single-cell RNA-seq



Read Counts

	Cell 1	Cell 2	...
Gene 1	18	0	
Gene 2	1010	506	
Gene 3	0	49	
Gene 4	22	0	
...			

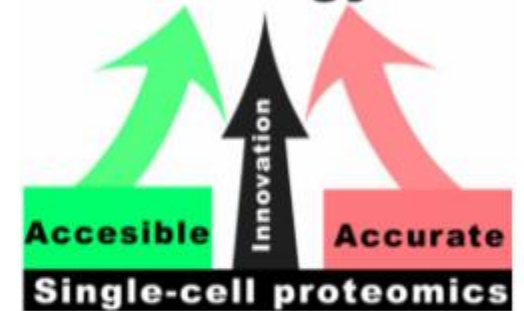
Compare gene expression profiles of single cells



translation

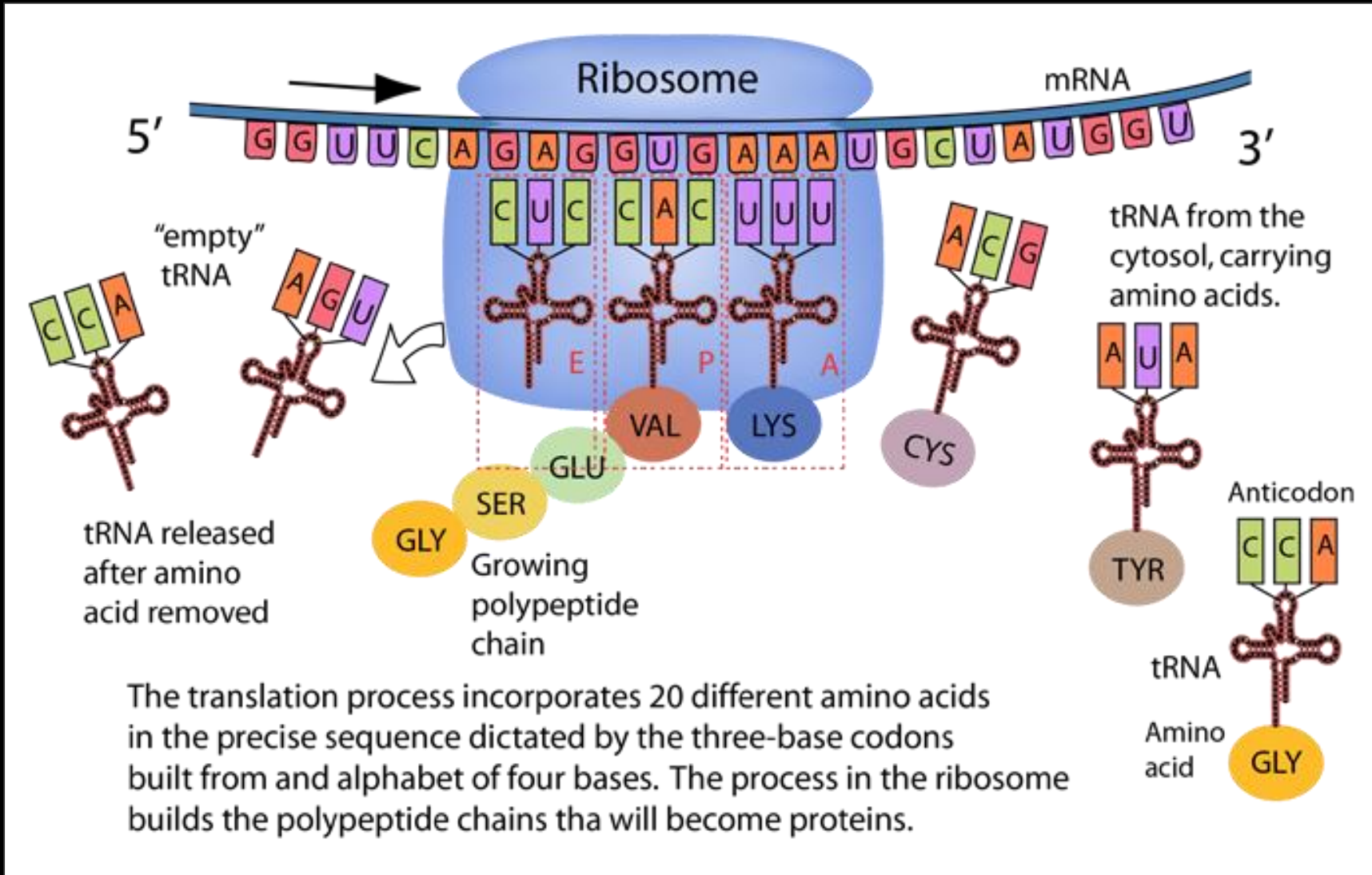
### Single-cell proteomics

#### Single-cell biology



	Cell1	Cell2	Cell3	...
Protein1				
Protein2				
Protein3				
Protein4				
...				

# Translation

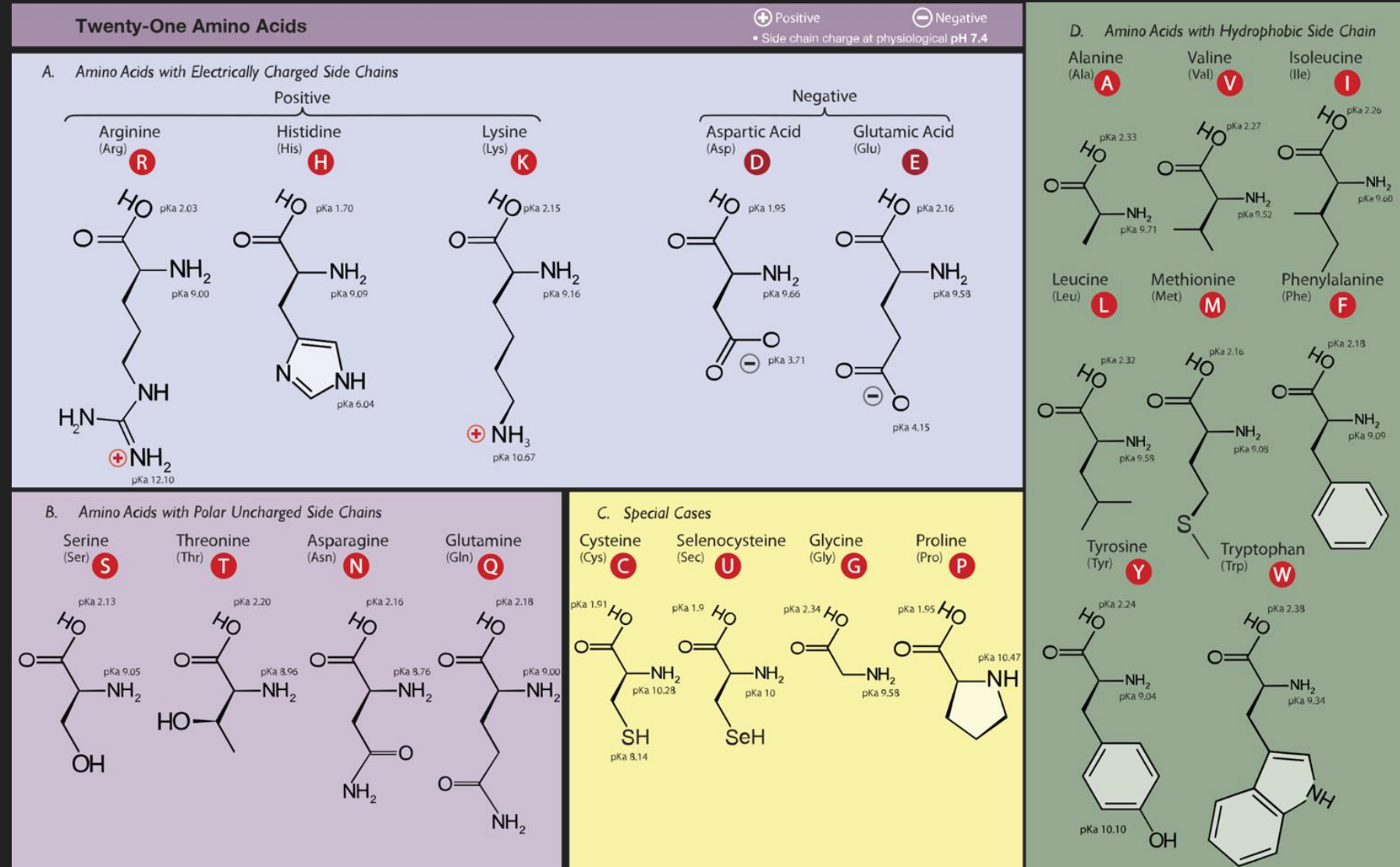


# Protein composition

Amino acids have specific properties such as

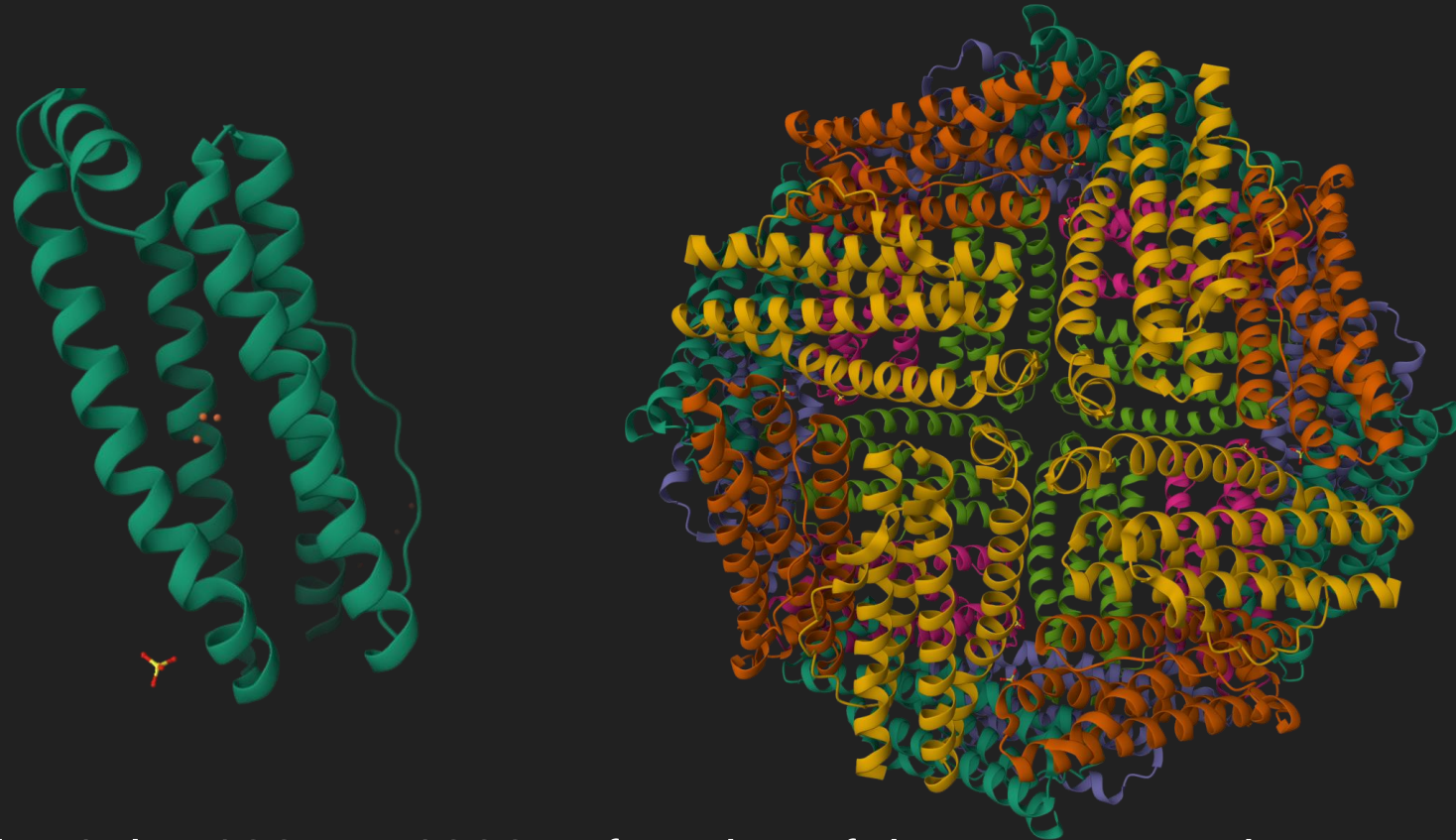
- charged
- uncharged
- hydrophobic

Those can create bonds between peptides





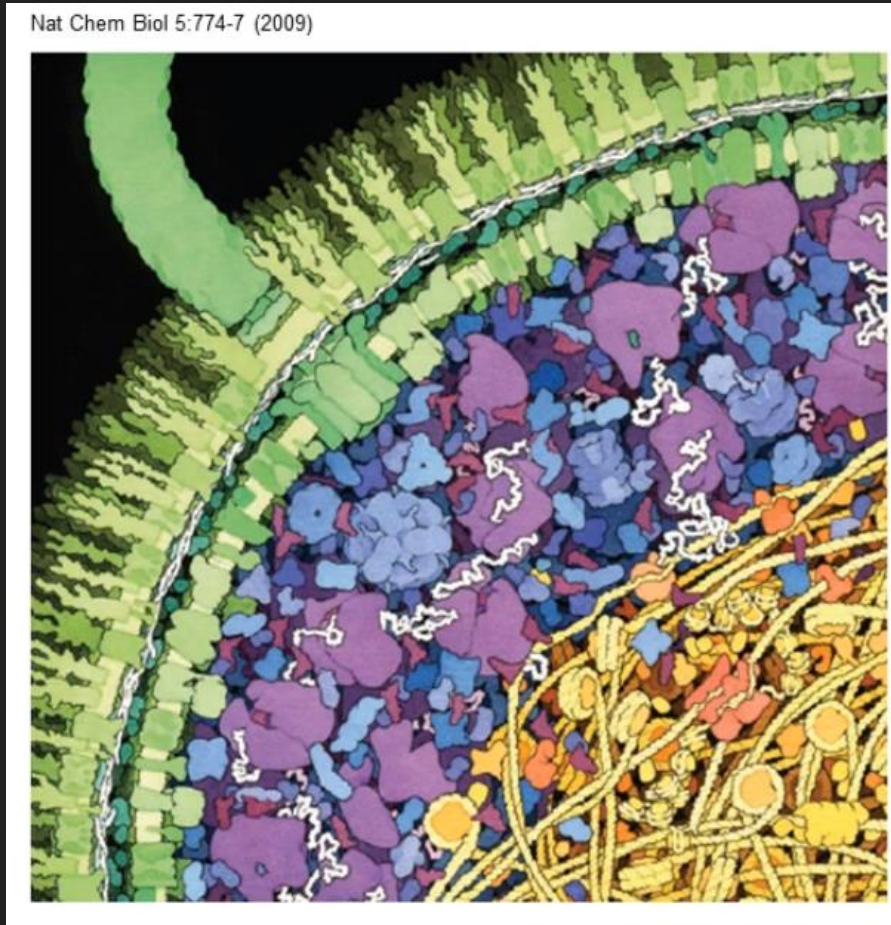
# A lot of proteins!



Not 1, but 100s or 1000s of copies of the same protein

A Marchetti *et al.* *Nature* **000**, 1-4 (2008)  
doi:10.1038/nature07539

# Protein/peptide isolation



Challenge:

isolate proteins to detect the distinct chemical structures

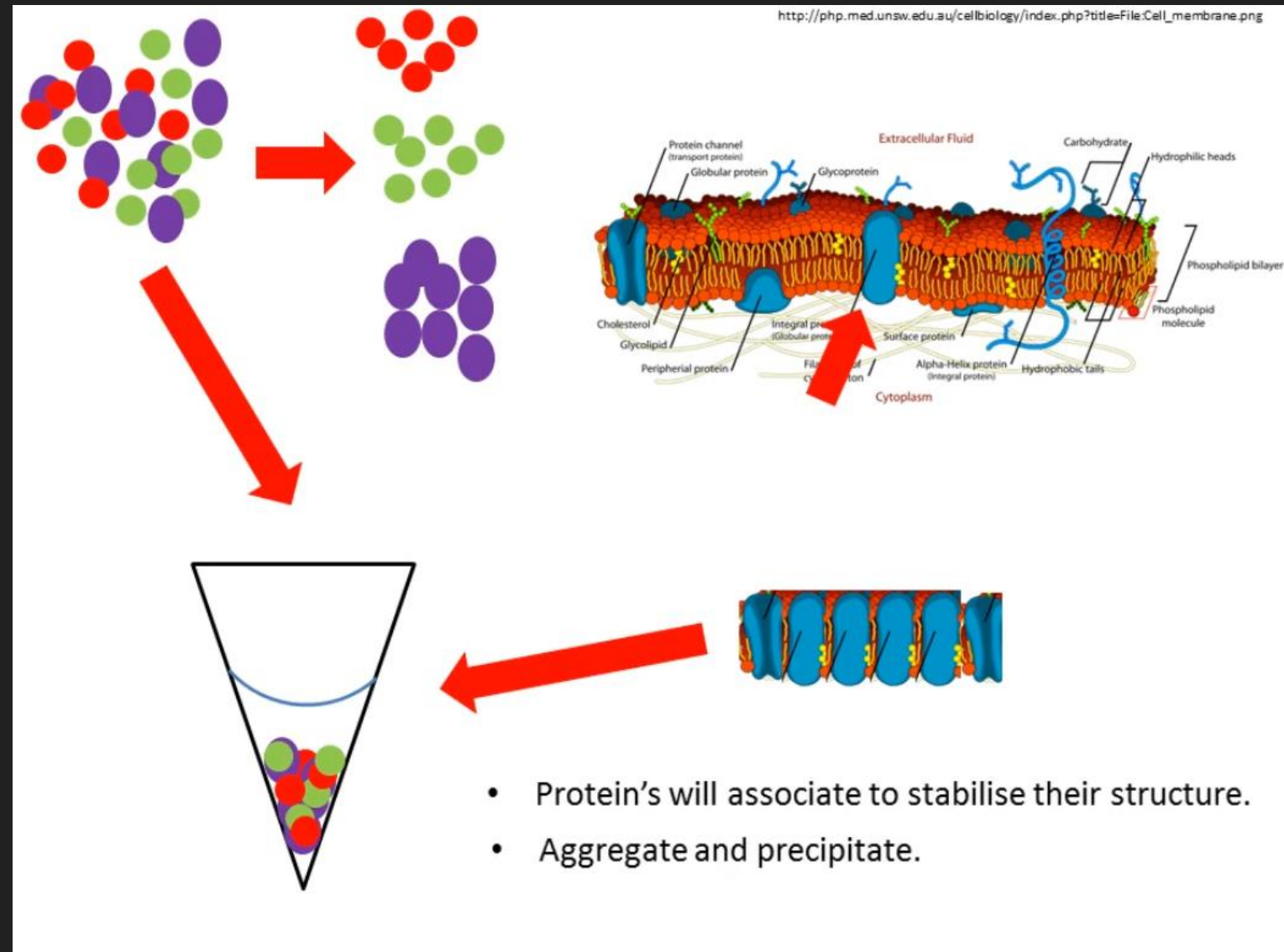
Note: there is no PCR here



# Protein/peptide isolation

## Fluidic isolation

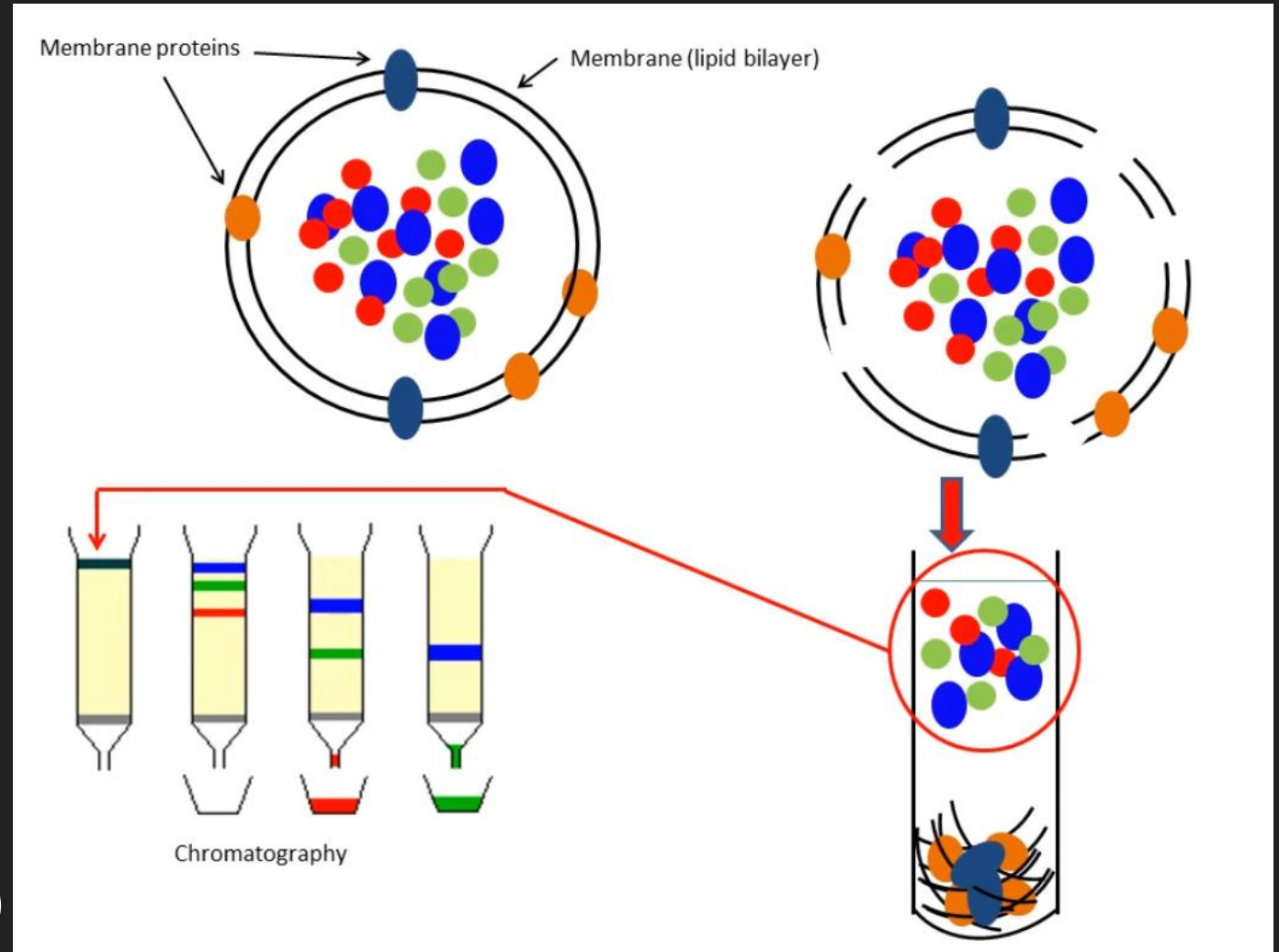
- recovers some peptides
- new associations to stabilize peptides (e.g. hydrophobic bonds)
- heavy bonds will be lost by precipitation



# Protein/peptide isolation

## Fluidic isolation

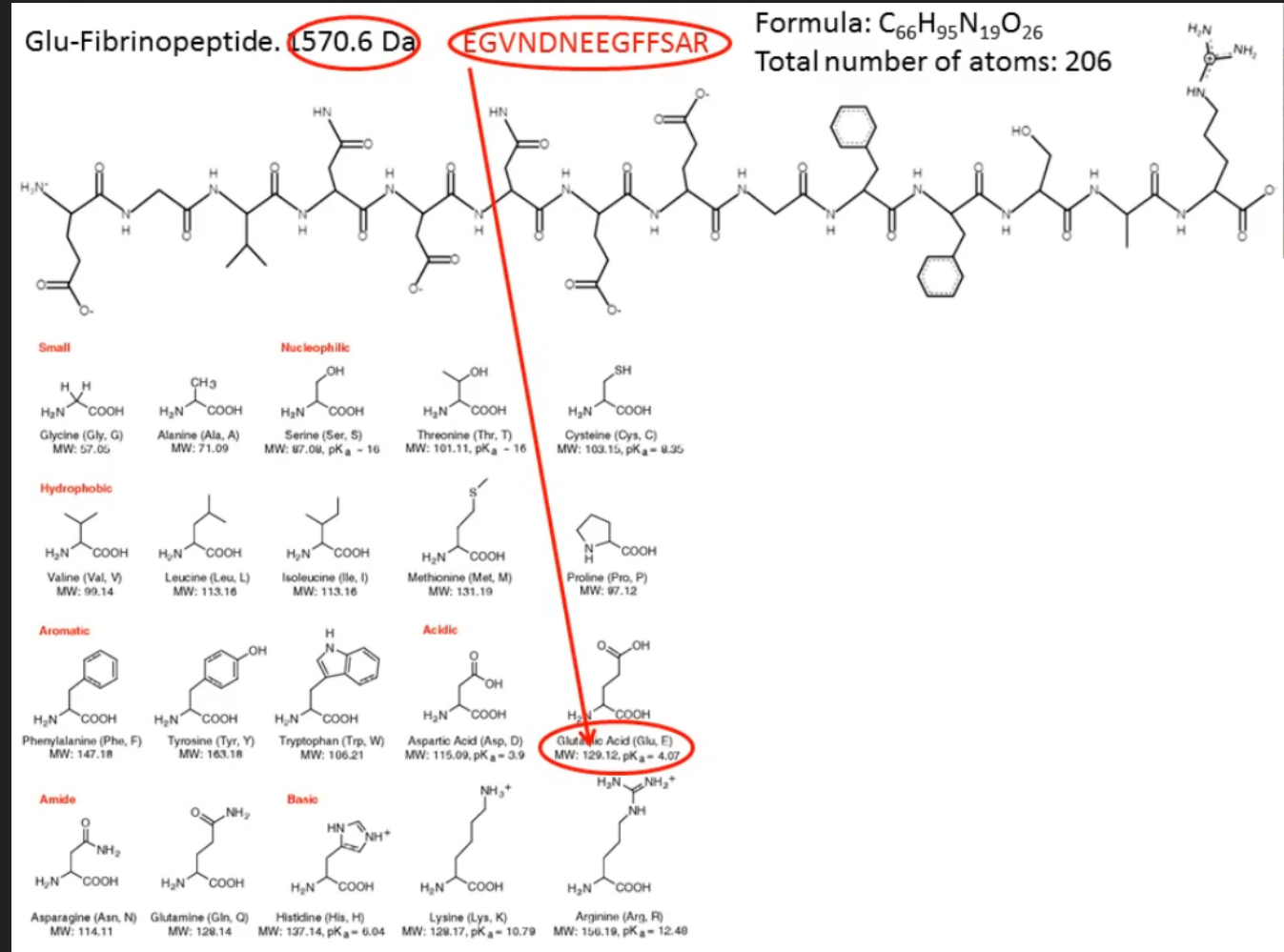
- centrifugation recovers more peptides
- a chromatograph can separate them
- specific liquid solutions reduce the precipitation loss further by avoiding bonds (e.g. hydrophobic)



# Mass Spectrometry rationale

## Mass is important

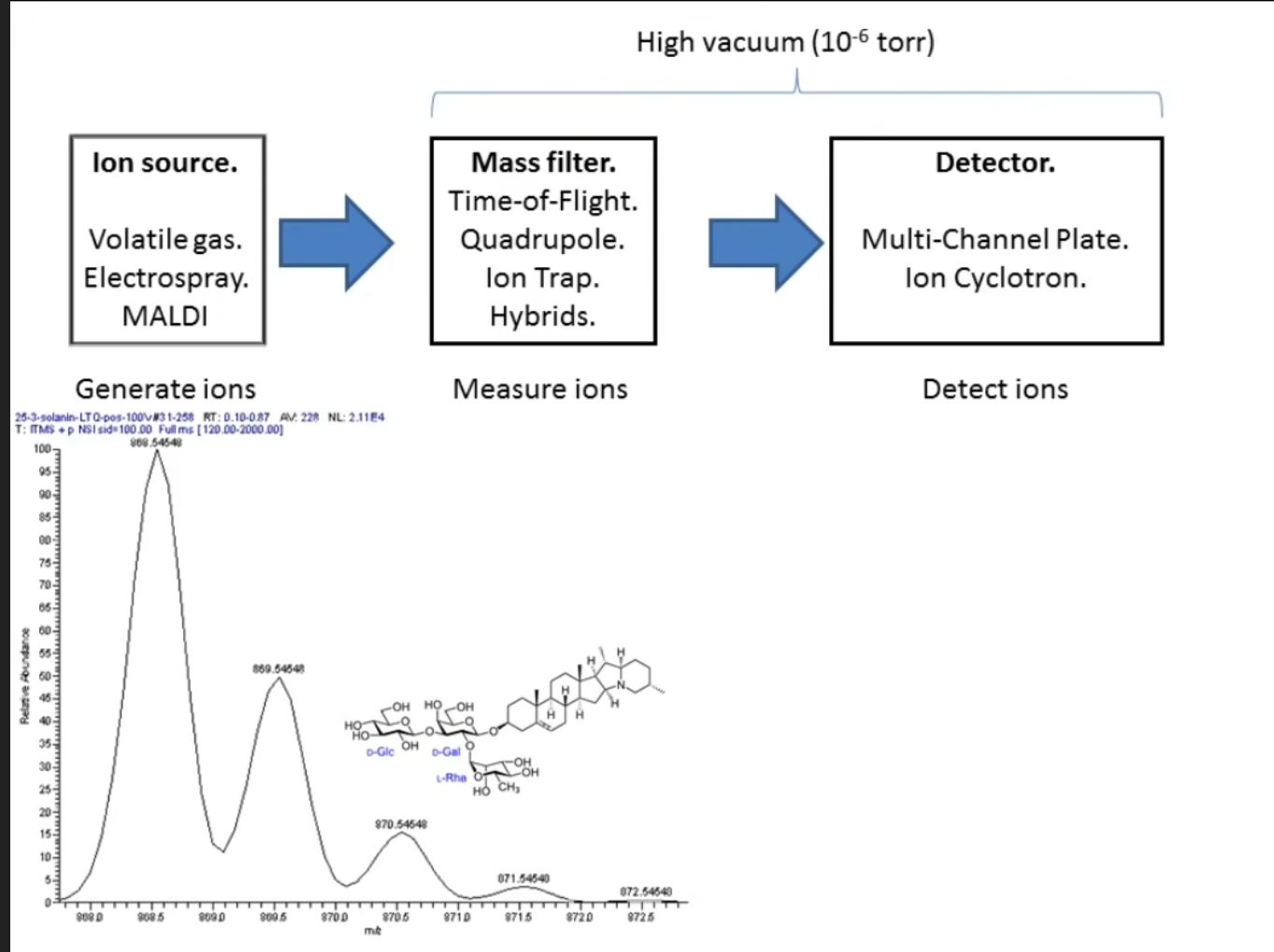
- elements in the periodic table have unique mass
- their compounds also do, including peptides
- We cannot calculate the mass of peptides when we do not know the sequence



Matt Padula

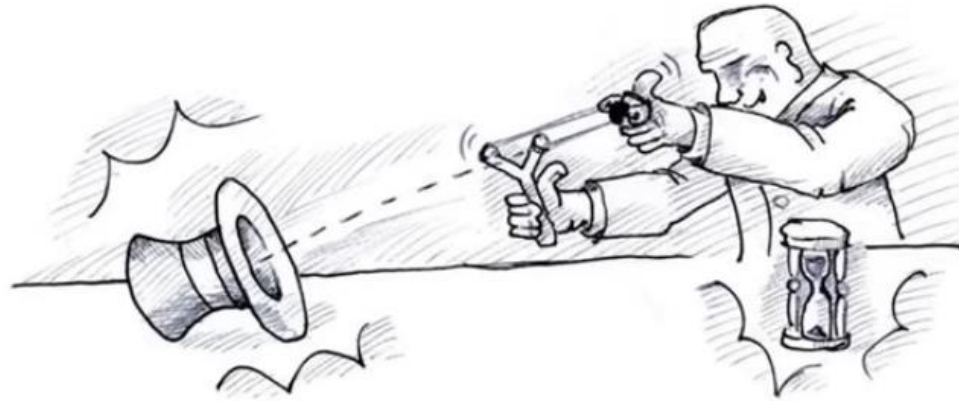
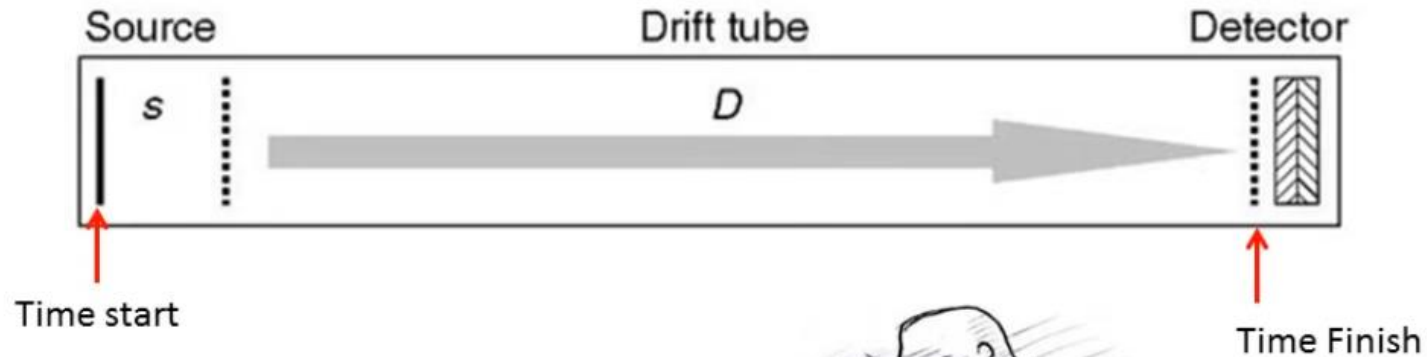
# Mass Spectrometry

- We charge our particles
- We make them go through a dedicated device (e.g. a Time Of Flight device)
- We measure their mass



# Mass Spectrometry

- Time of Flight: The time it takes an ion to go from ion source to detector is directly related to its mass.
  - Heavier particles reach lower speeds.



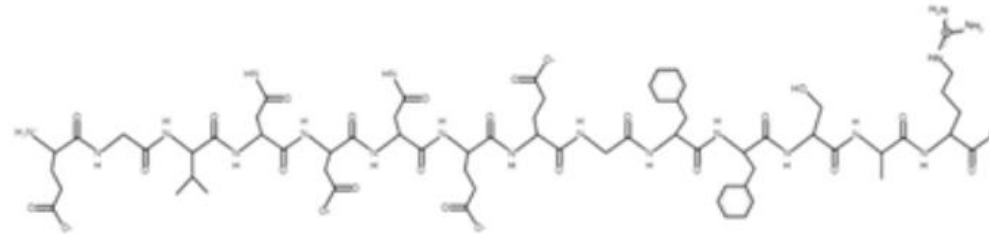
$$\text{velocity} = \sqrt{\frac{2 \times \text{energy}}{\text{mass}}}$$

$$\text{Flight\_time} = \frac{\text{drift\_length}}{\text{velocity}} = \text{drift\_length} \times \sqrt{\frac{\text{mass}}{2 \times \text{energy}}}$$



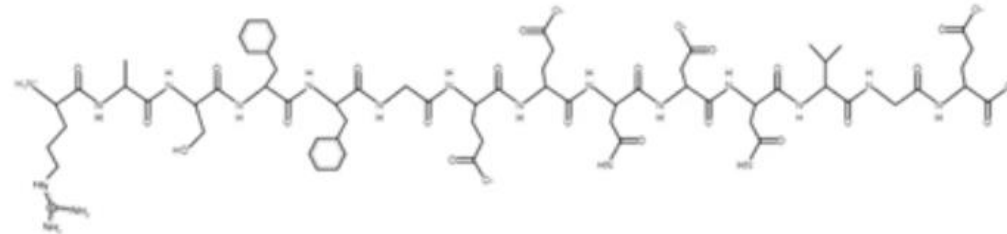
# Mass Spectrometry

- But, in a more complex mixture of peptides from different protein isoforms, isobaric peptides exist.



EGVNDNEEGFFSAR

1570 Da

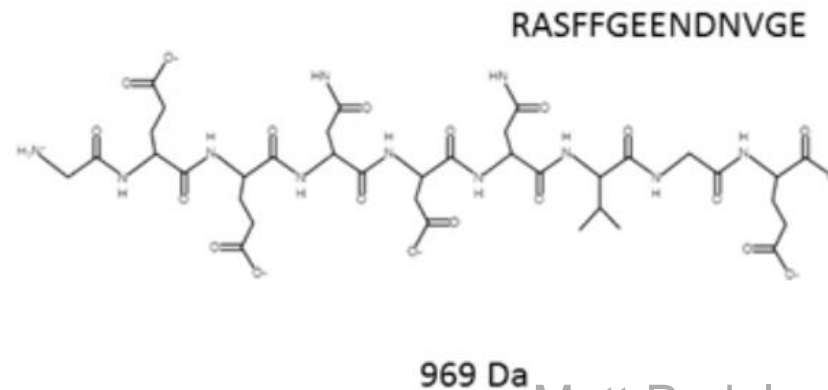
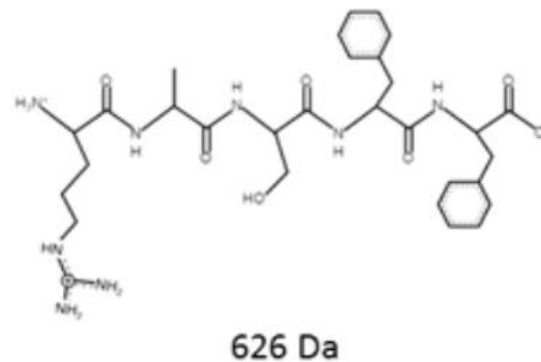
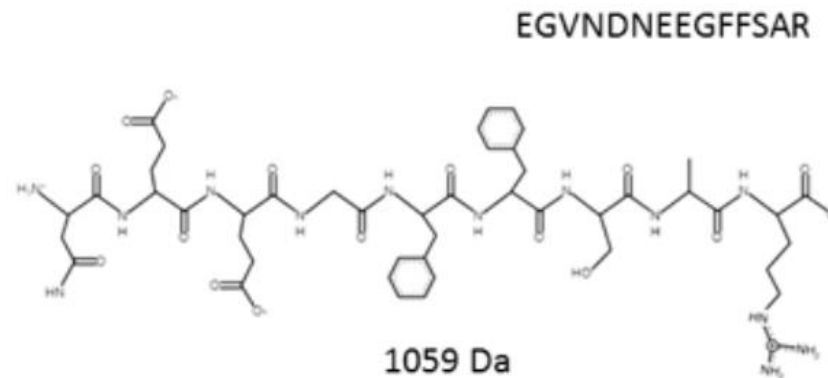
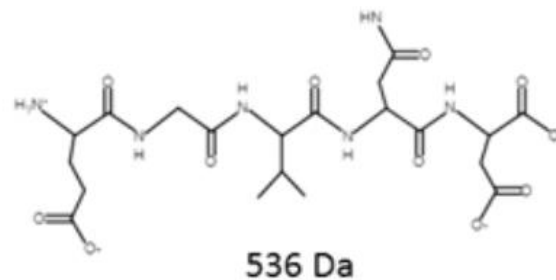


RASFFGEENDNVGE

- Same mass, but different structure.
- Measuring intact mass doesn't reveal isomers.

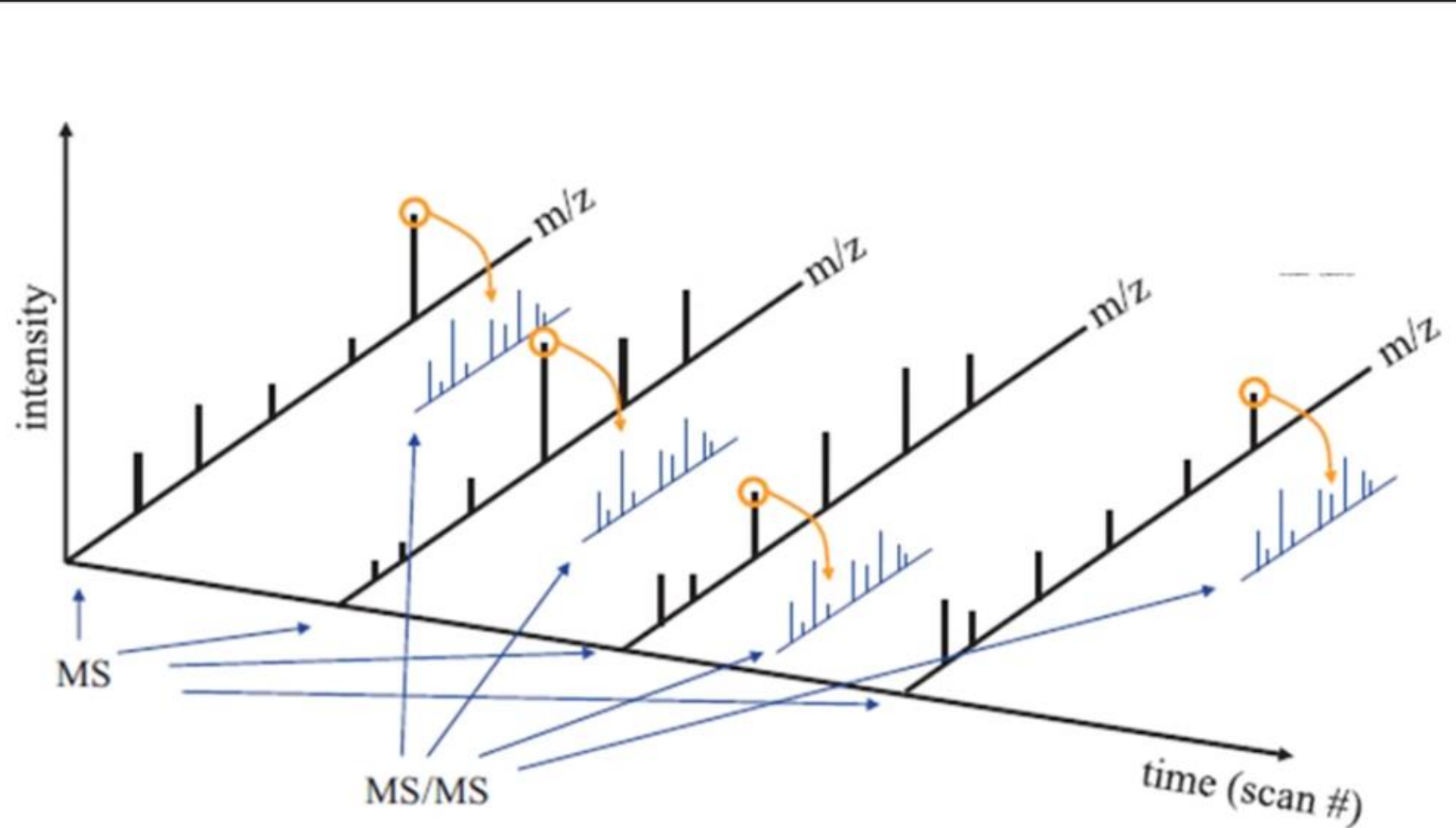
# Mass Spectrometry

- Break molecule into smaller pieces and measure their masses.



# Mass Spectrometry

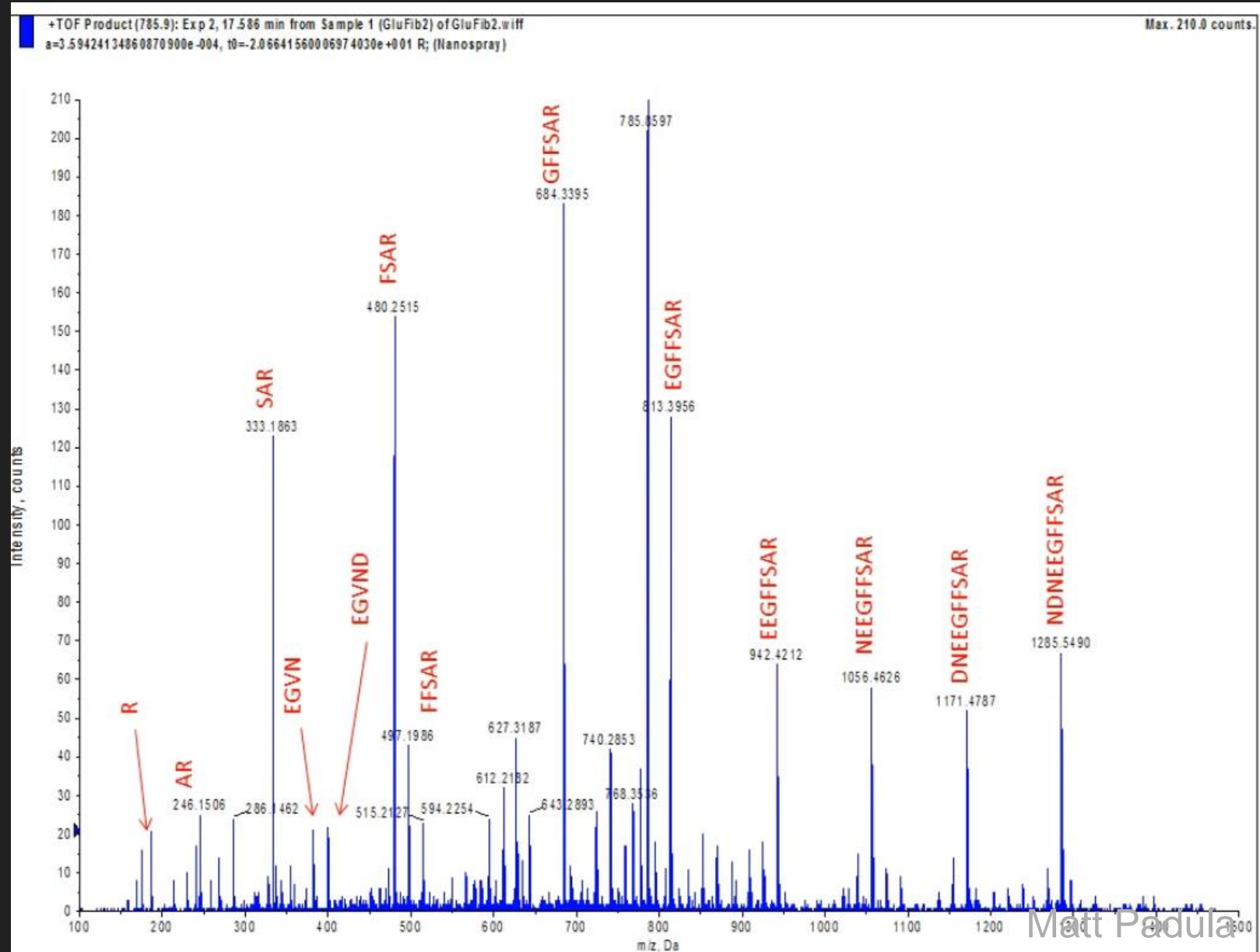
The MS scans continuously for intact peptide ions, selecting one of those ions, fragmenting it, and measuring the mass of each fragment



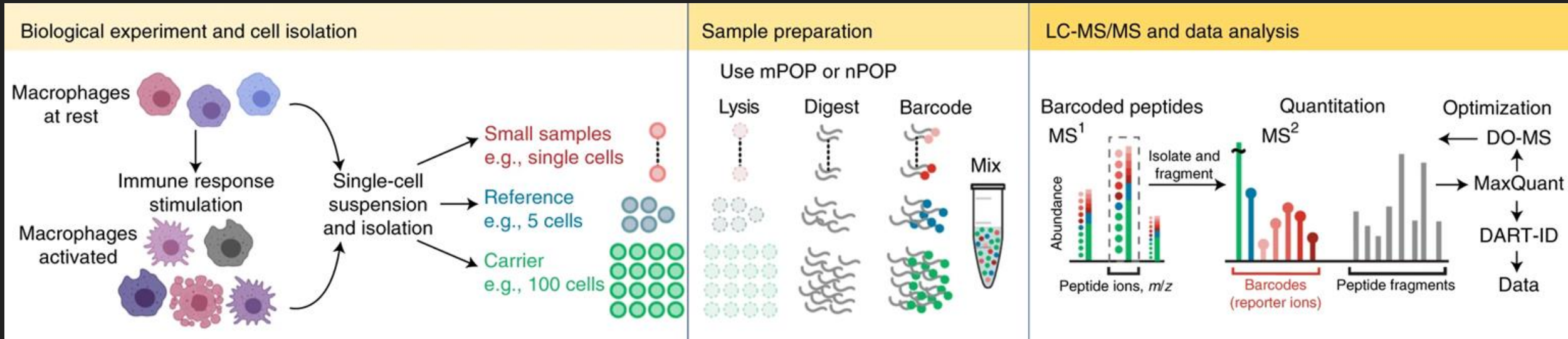
# Mass Spectrometry

The compound sequence is not known, but can be built back by putting together peaks which difference is the mass of an amino acid.

Various software do that (e.g. MaxQuant)



# Single Cell Mass Spectrometry SCoPE2 technology



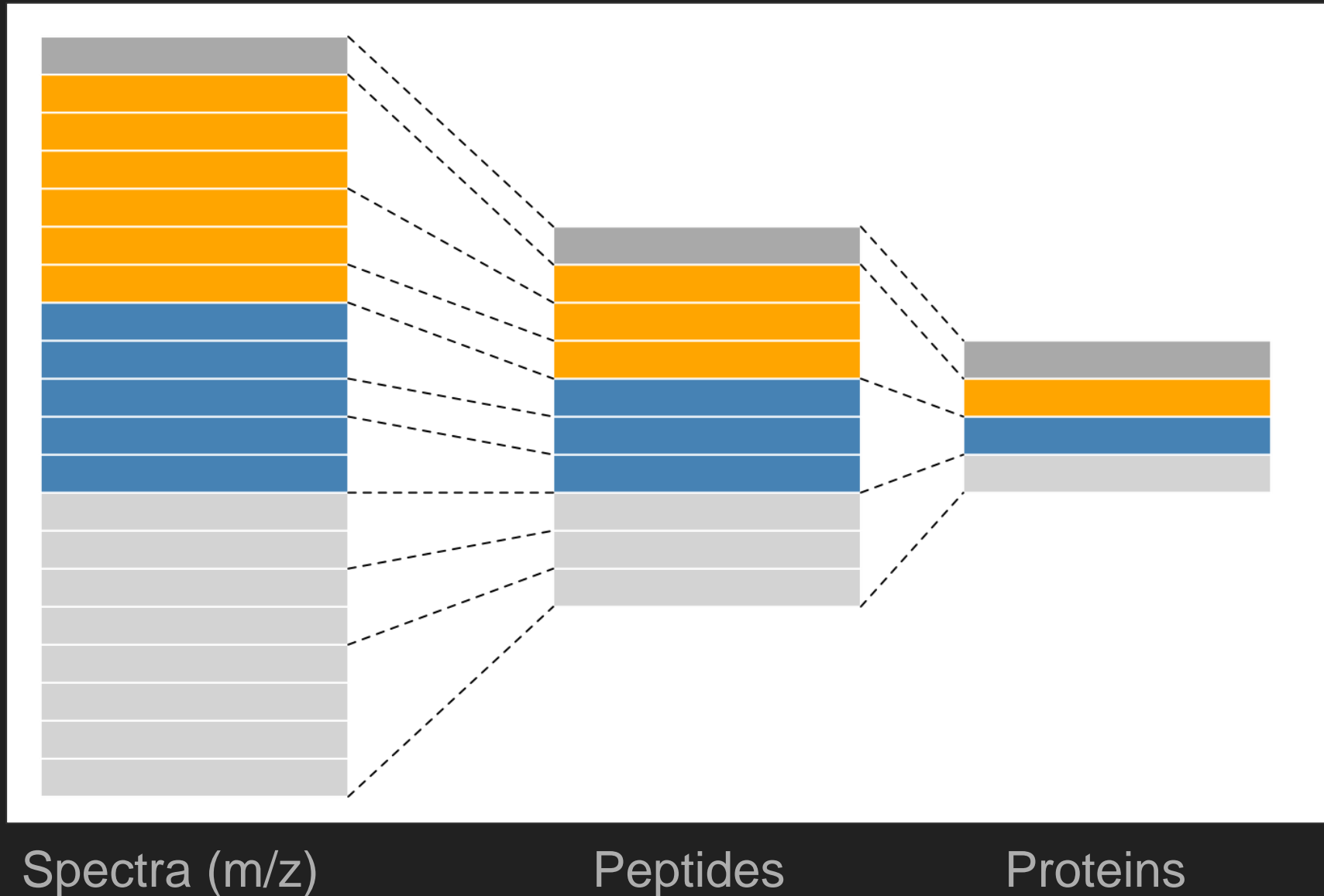
Cells isolation and spiking of carrier bulk of peptides, which increases the amount of peptides. The carrier will be later removed under demultiplexing.

Barcoding with TMT (Tandem Mass Tags) to recognize cells and normalize intensity signals. MaxQuant + DO-MS, DART-ID for peptides and protein ID.



# Single Cell Mass Spectrometry

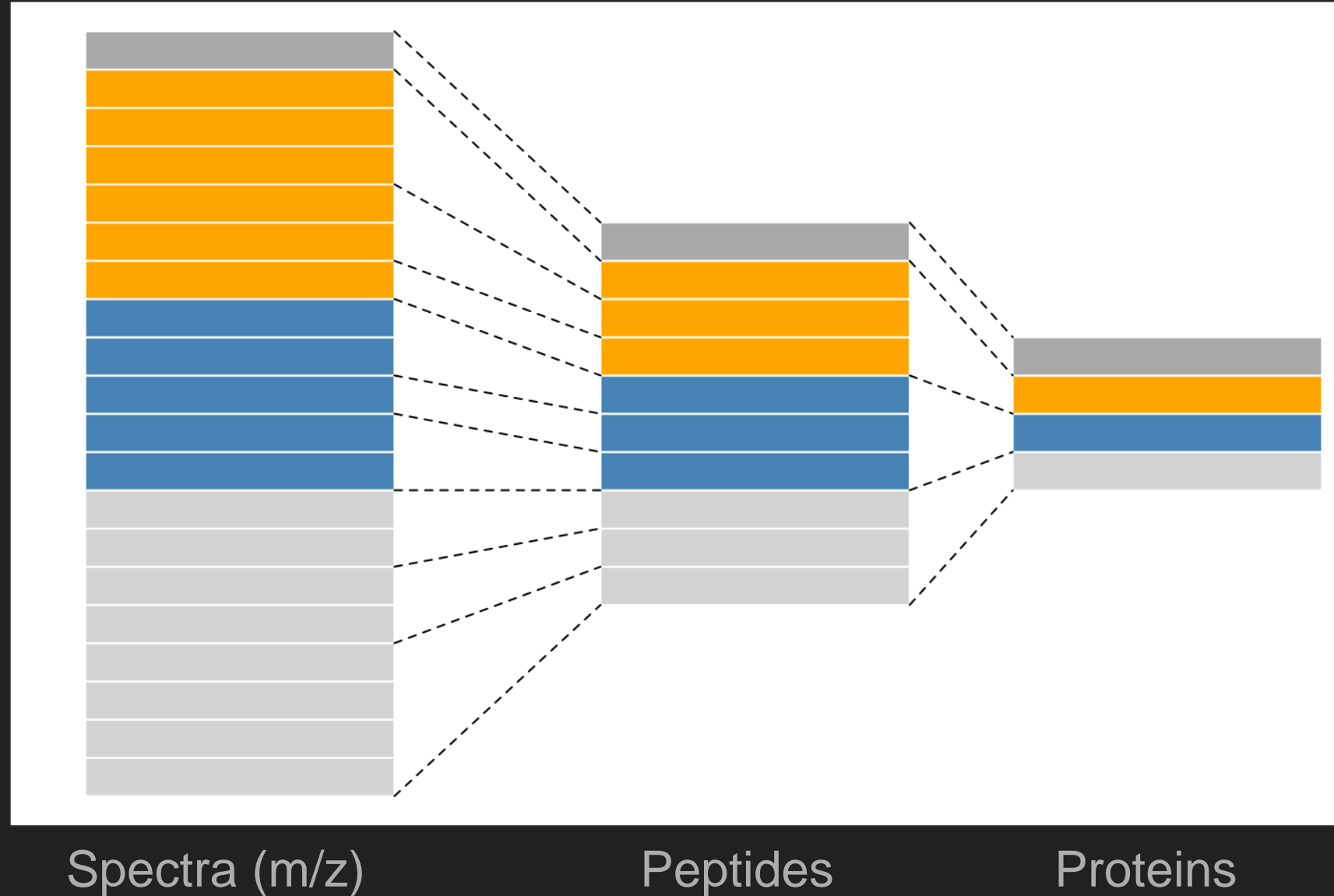
## Multilayered data format



# Single Cell Mass Spectrometry

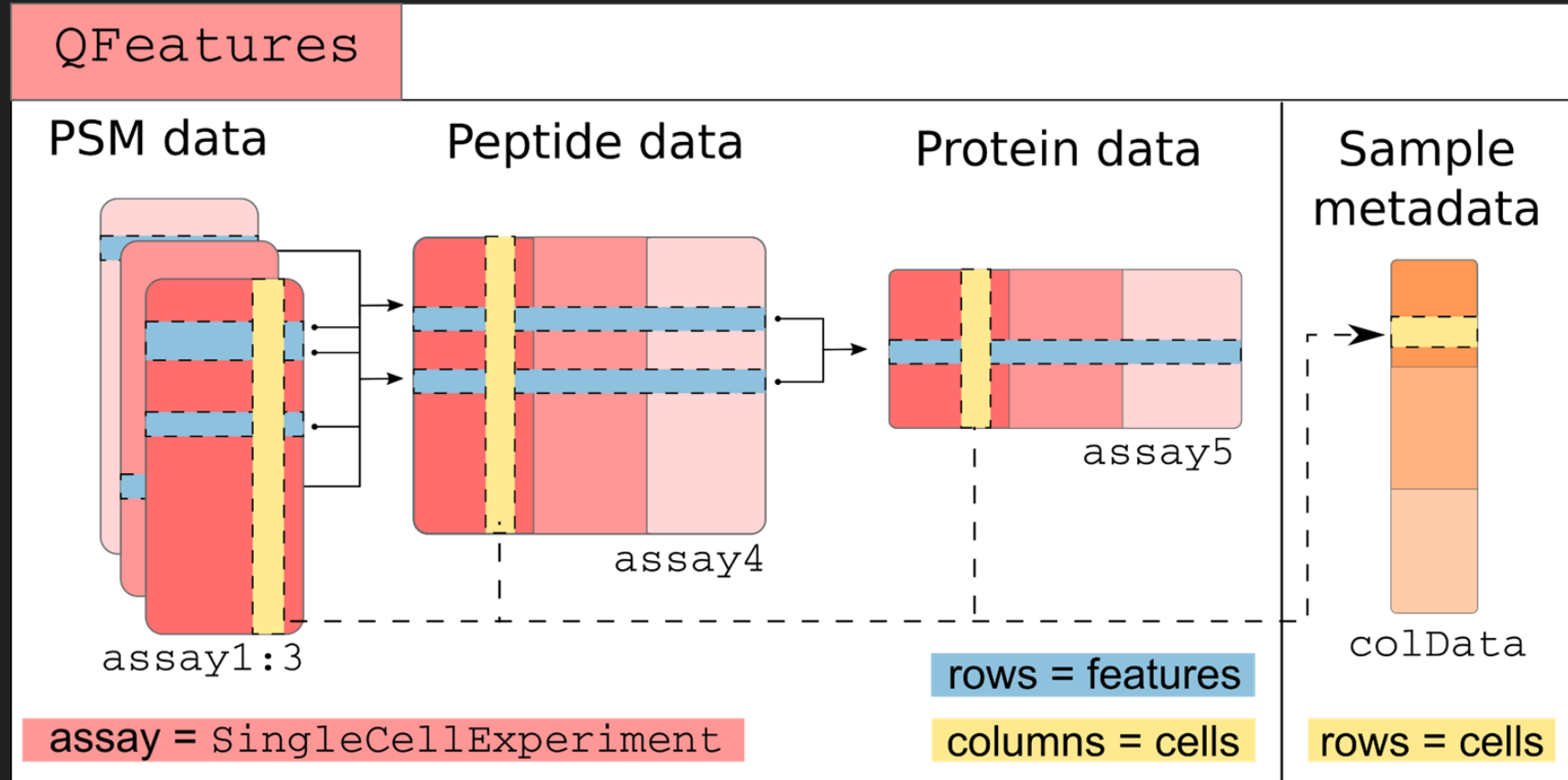
## Multilayered data format

Needs for storing  
all steps from raw  
output of MS to proteins



# Single Cell Mass Spectrometry

## Multilayered data format



QFeatures: reuses the **singlecellexperiment** format to include all the protein information

# Single Cell Mass Spectrometry References/Links

[Slavov Lab](#) (SCoPE technology)

SCoPE [Paper](#)

SCoPE2 [Paper](#) and [Webpage](#)

SPDB Database [Paper](#) and [Webpage](#)

WetLab-side [review](#)

DryLab-side [review](#)

# Tutorial and open coding

At our home <https://abc.au.dk/Documentation> you can find

- SCP tutorial on single cell proteomics data
- Previous tutorials + 1 conference workshop

Or you can code and ask for help or generic coding/bioinf/data science questions



And have cake