

# Biophysical Chemistry

---

Justin Benesch

Jonathan Doye

Mark Wallace

# Lecture 7 - Biomolecular Mass Spectrometry

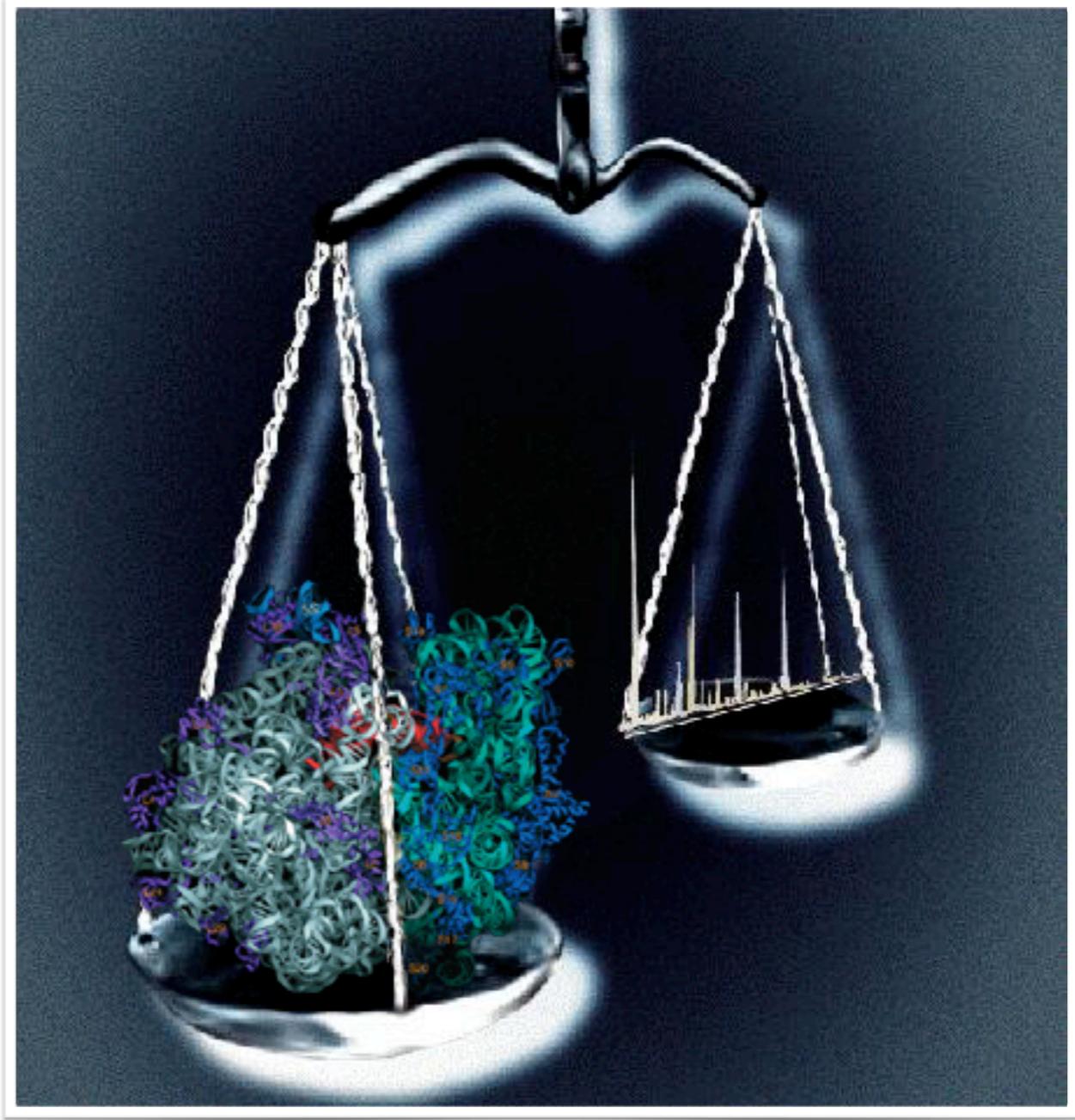
---

- **Principles of mass spectrometry**

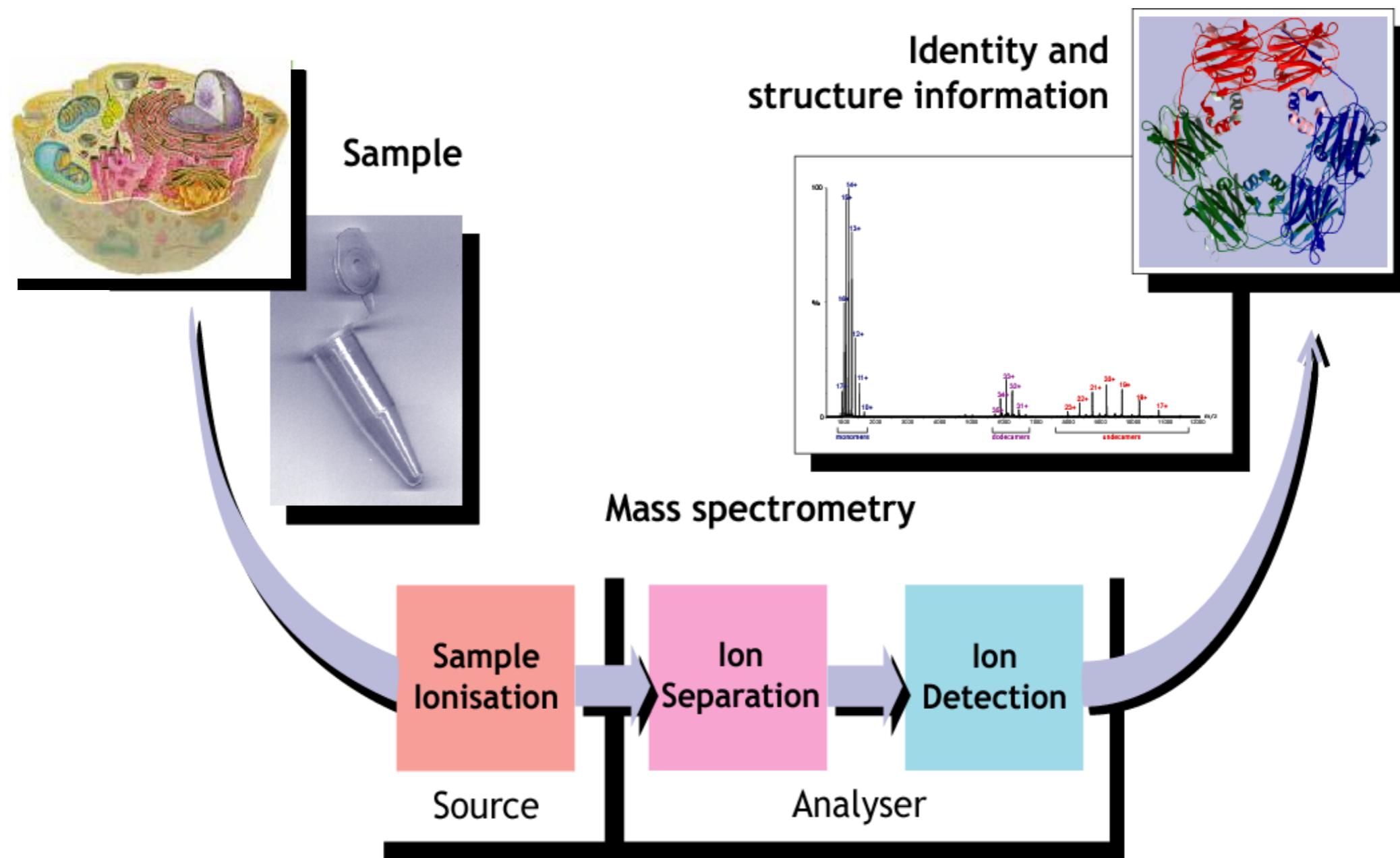
- Methods for ion formation
- Methods for ion separation
- Methods for ion activation

- **Application to biomolecules**

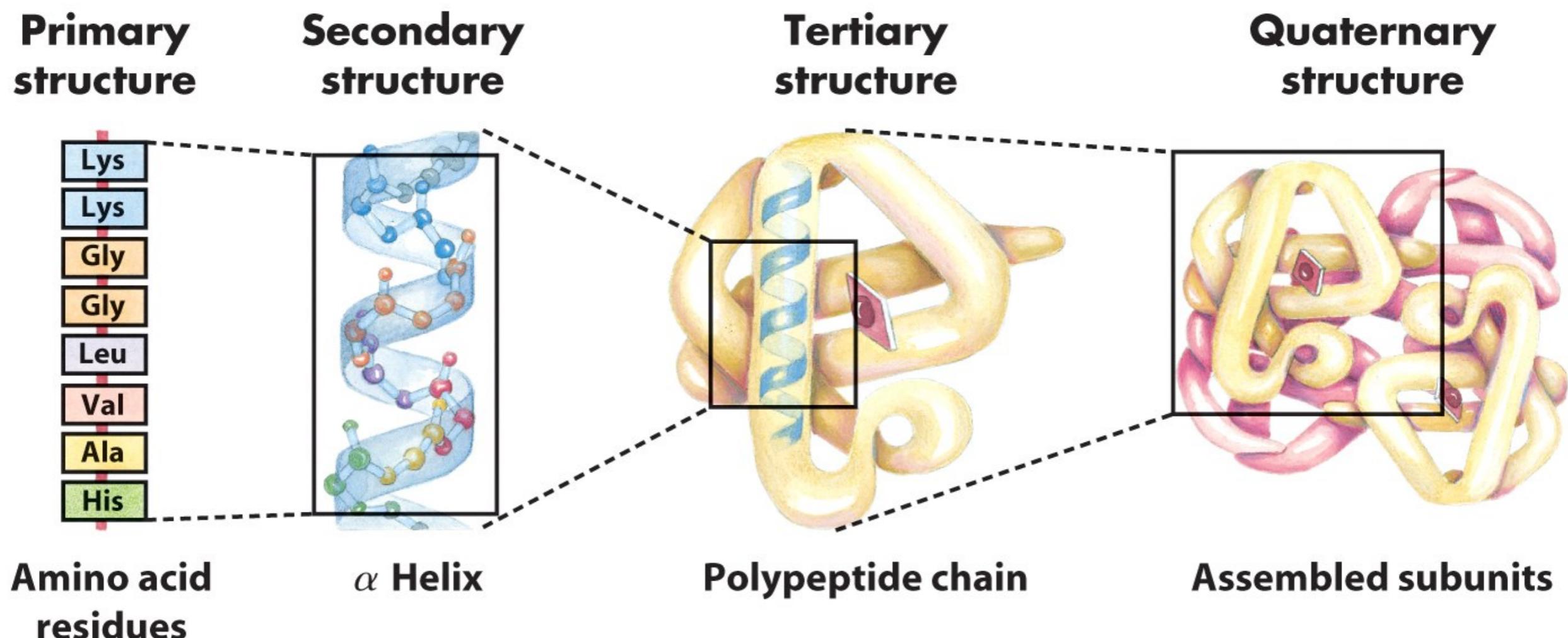
- Mass measurement
- Peptide sequencing
- Labelling technologies



# Mass spectrometry overview



# Levels of protein structure



- Mass spectrometry can inform on all levels of protein structure

# 'Soft' Ionisation

---

- From molecules in solution to ions in vacuum
- Established ionisation techniques resulted in covalent fragmentation of molecules
- Soft ionisation techniques are ionisation of large ions without their fragmentation
- Current 'mass record' is >100 MDa!
- Soft ionisation techniques have revolutionised 'proteomics' the study of the proteins encoded by the genome

## Nobel Prize in Chemistry 2002

"for their development of soft desorption ionisation methods for mass spectrometric analyses of biological macromolecules"

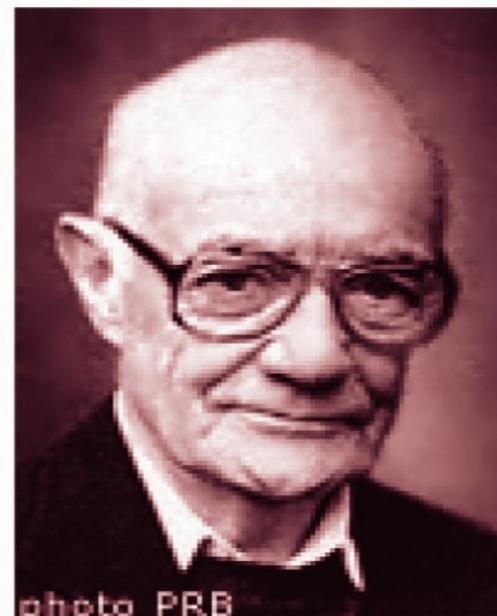


photo PRB

**John B. Fenn**

◐ 1/4 of the prize

USA



photo PRB

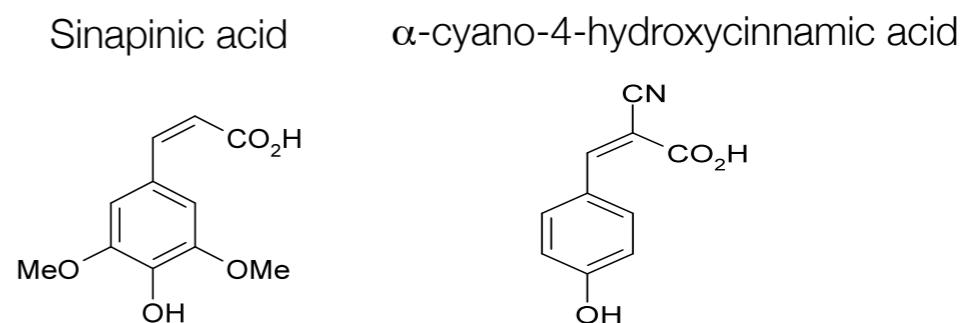
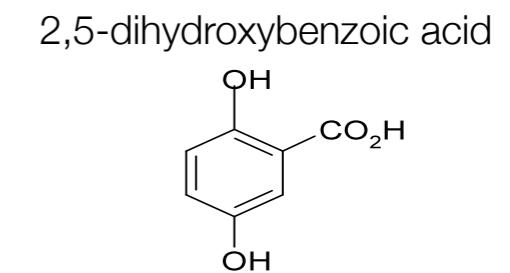
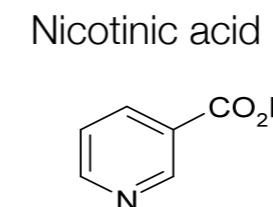
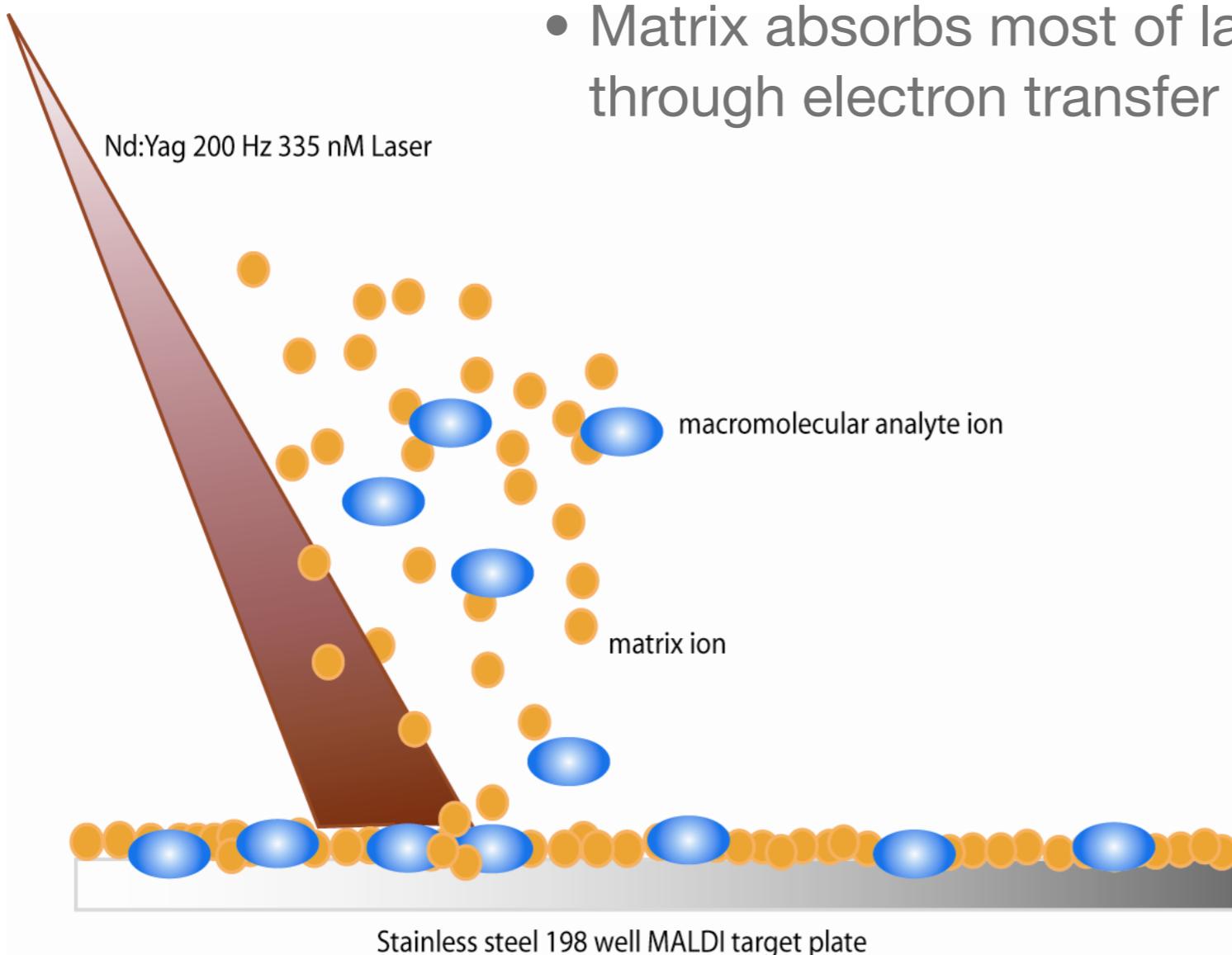
**Koichi Tanaka**

◐ 1/4 of the prize

Japan

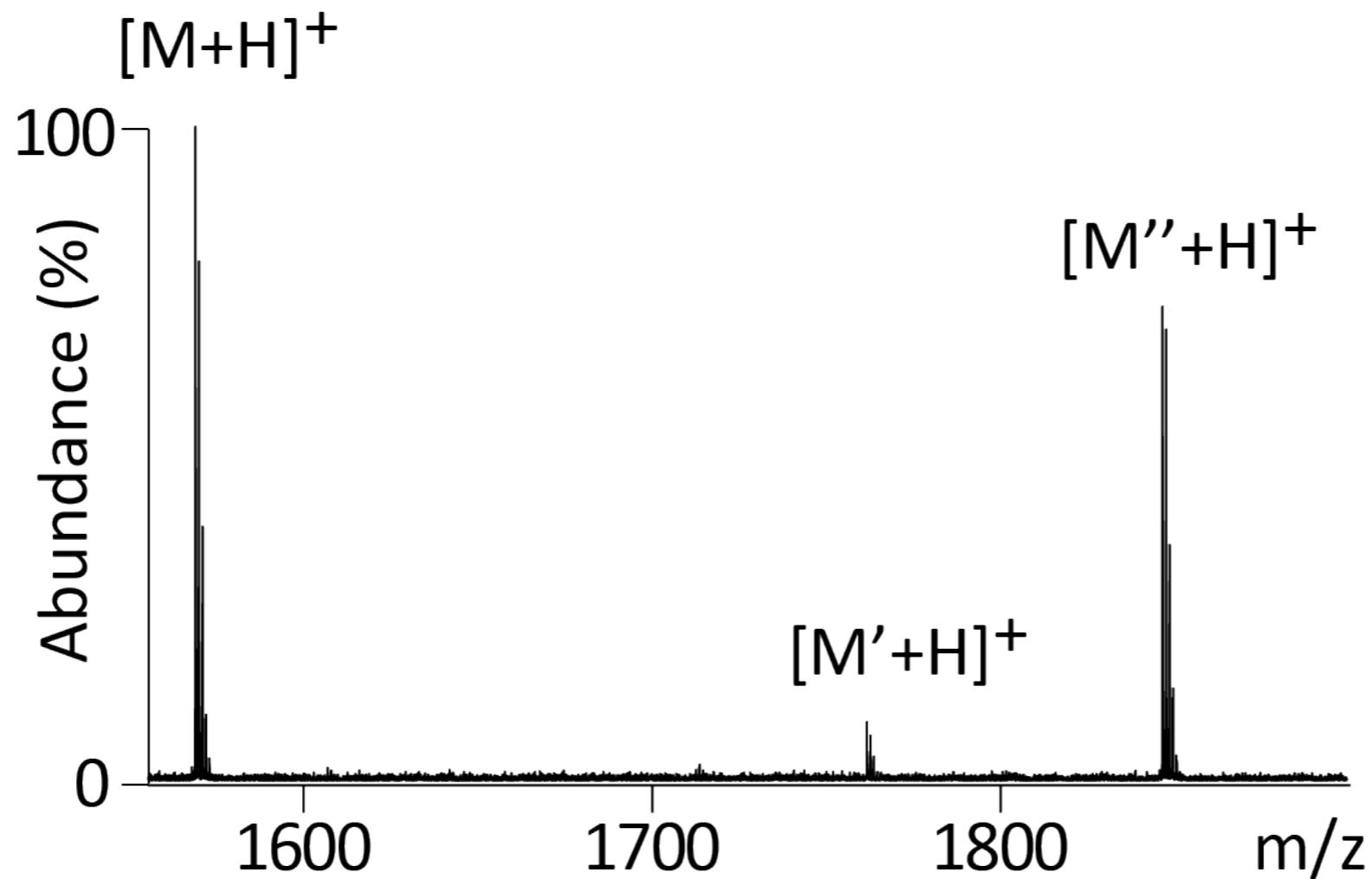
# Matrix-assisted laser desorption/ionisation (MALDI)

- Analyte is mixed with a large excess of matrix that absorbs at laser wavelength
- Matrix absorbs most of laser energy and aids ionisation through electron transfer and chemical processes



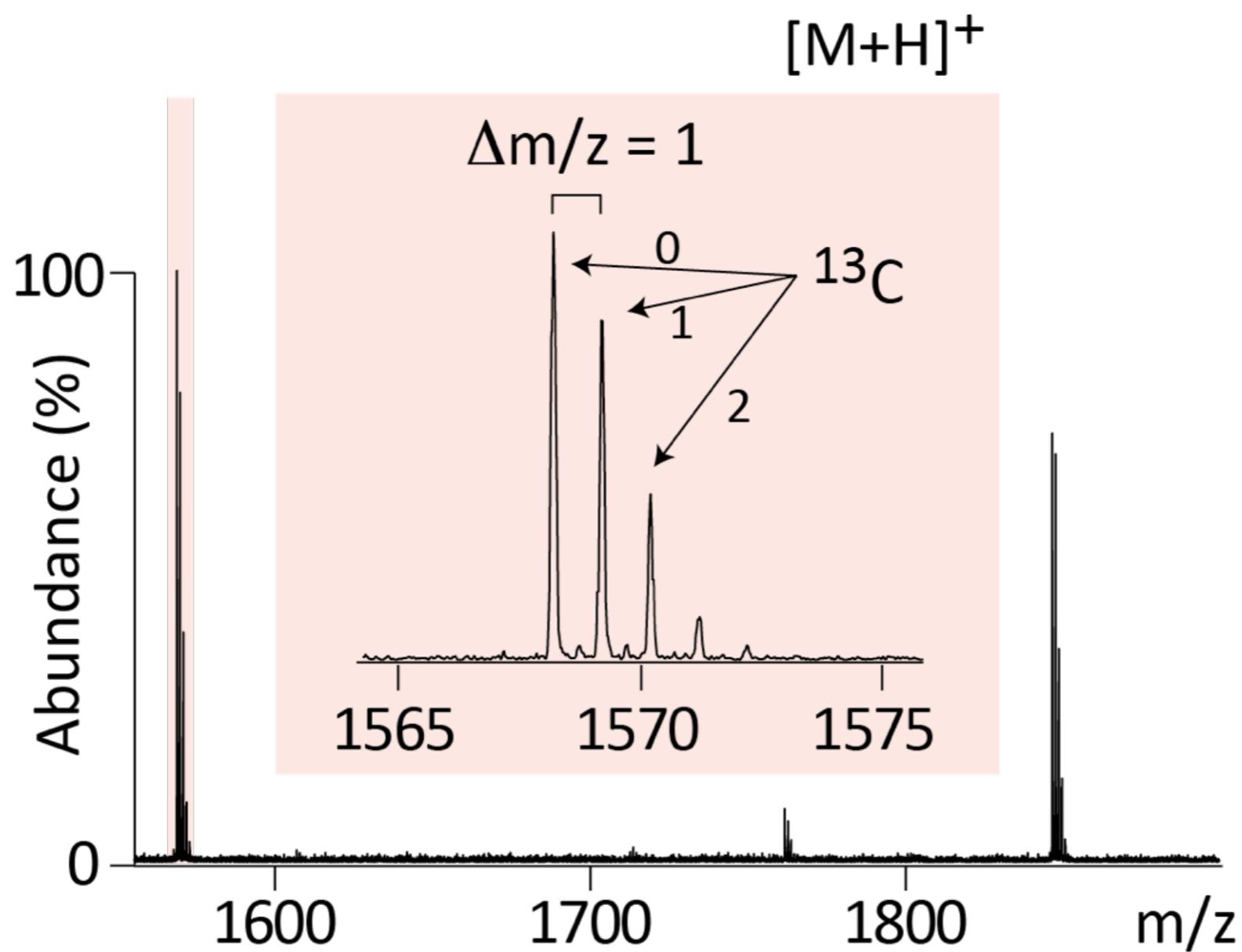
# MALDI spectrum of peptide mixture

---



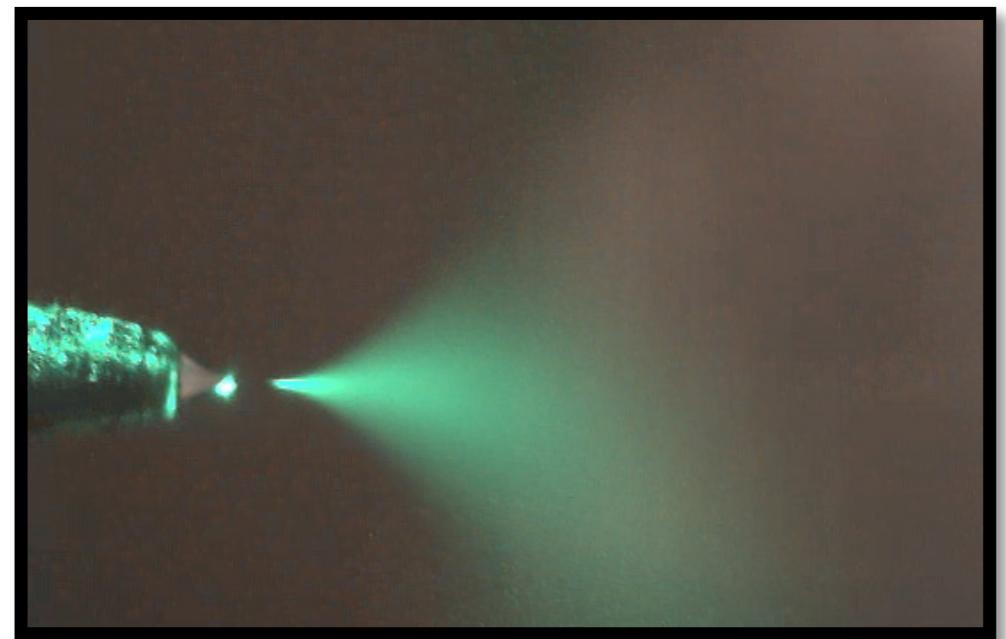
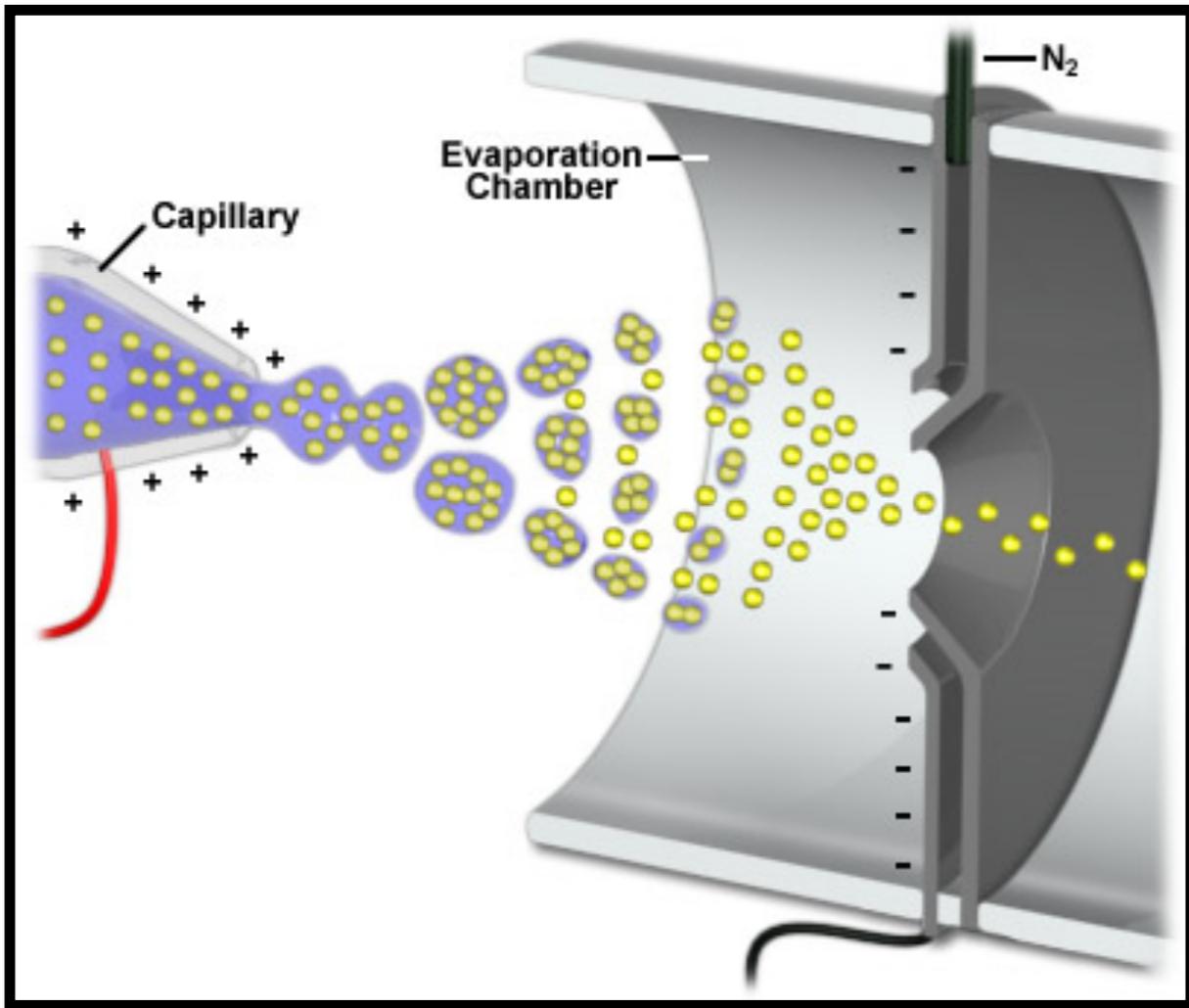
- Plot of signal intensity versus mass-to-charge (Da/e, or Thomson)
- Mixture of three peptides with mass  $M$ ,  $M'$ ,  $M''$
- MALDI leads to singly protonated ions

# Resolution of isotopes



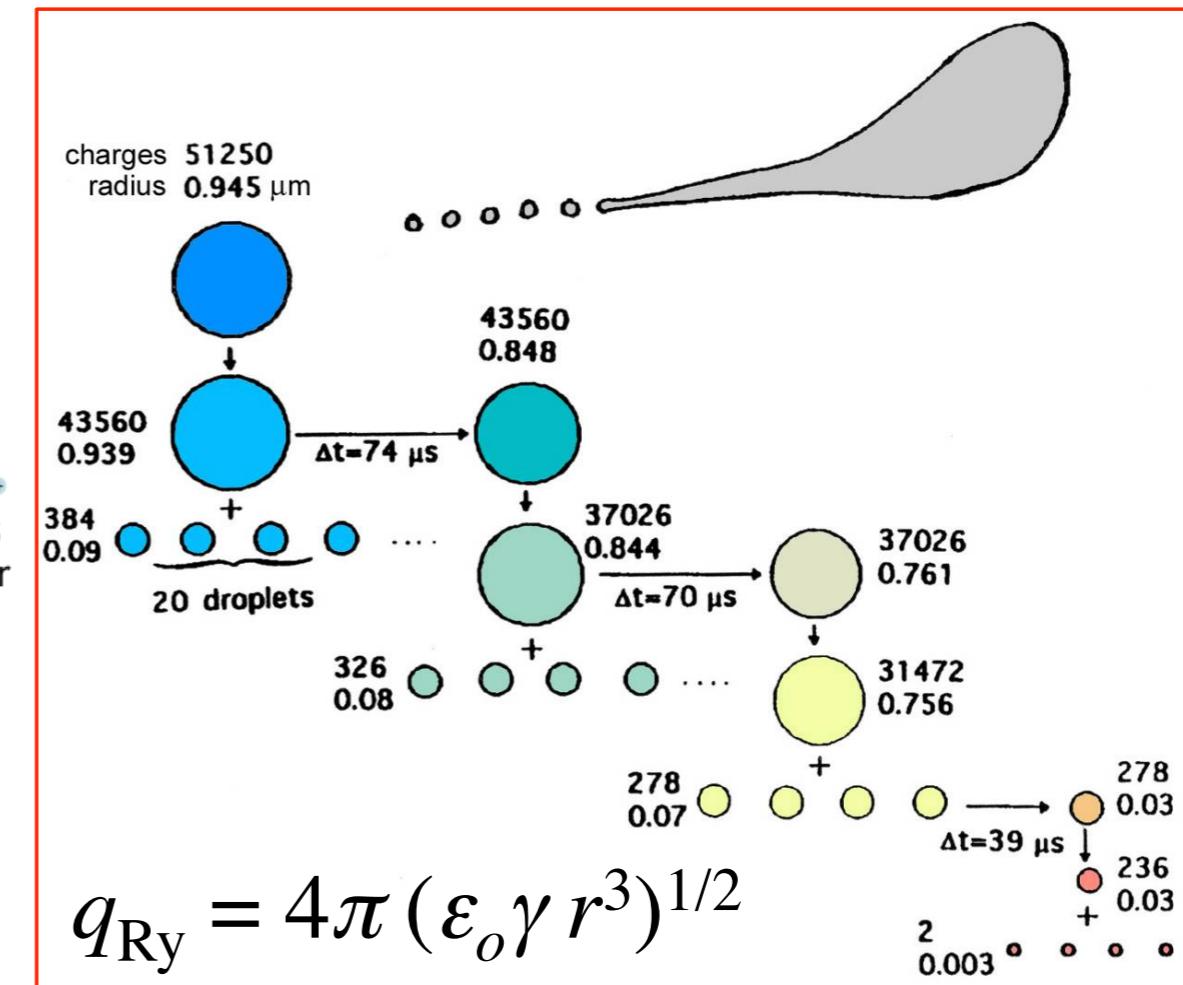
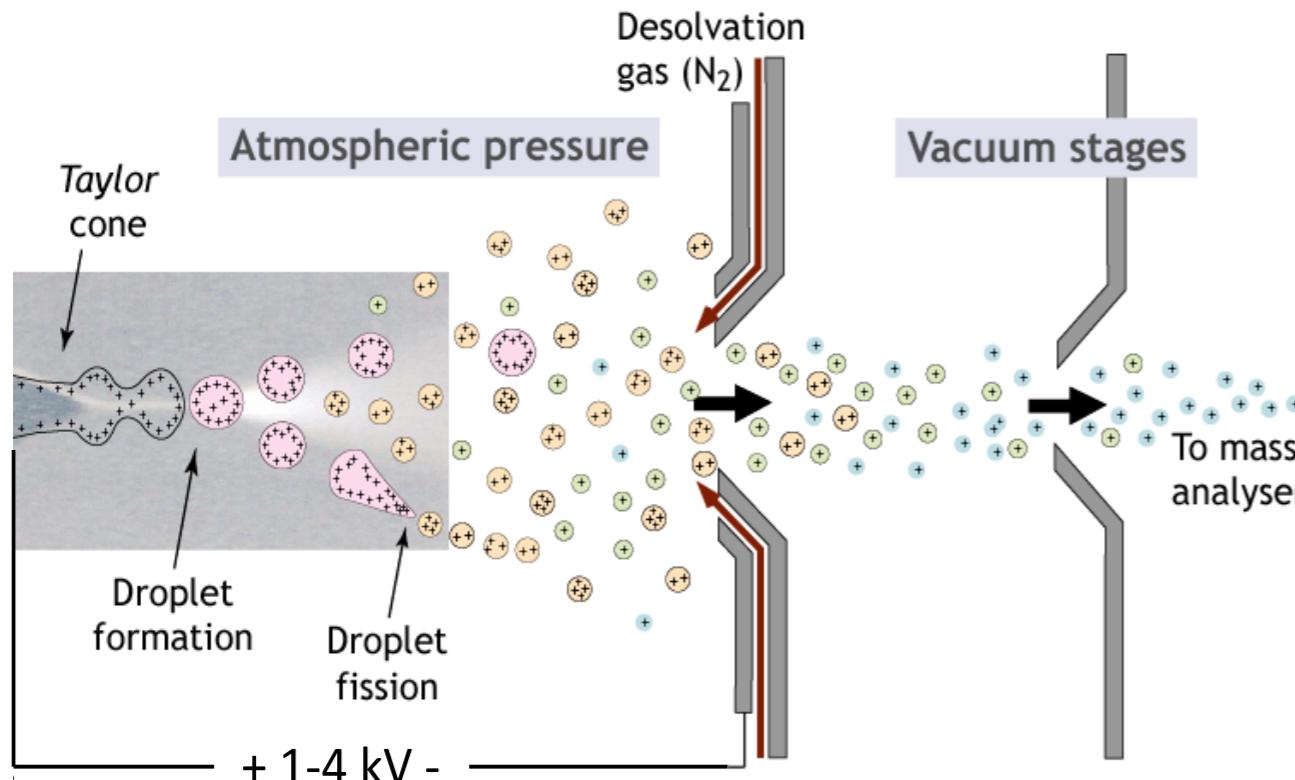
- Peaks reveal ‘fine structure’ due presence of  $^{13}\text{C}$  (natural abundance 1.1%)
- For a singly charged peptide,  $\Delta m/z = 1$ ; doubly charged peptide,  $\Delta m/z = 0.5\dots$

# Electrospray ionisation (ESI)



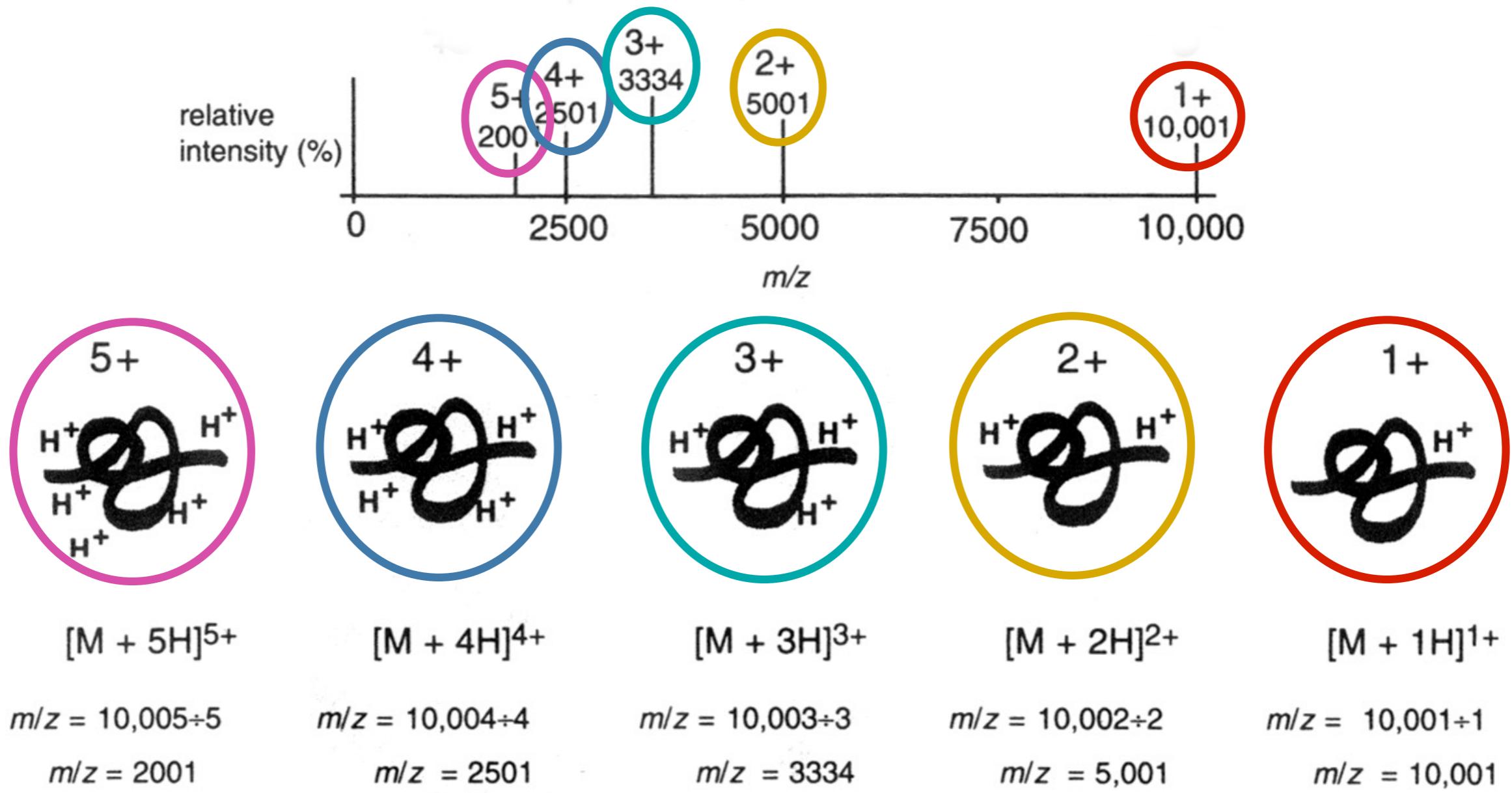
- Potential difference is applied between an emitter (needle) and a counter electrode
- Charged droplets are aerosolised and evaporation leads to ionised contents

# Mechanism of ion generation during ESI



- Droplets shrink by solvent evaporation and fissure at the Rayleigh Limit
- Some small ions are evaporated from droplet directly (ion evaporation model)
- Process proceeds until solvent is depleted and residual charge resides on contents (charged residue model, holds for proteins)

# ESI mass spectrum of single protein



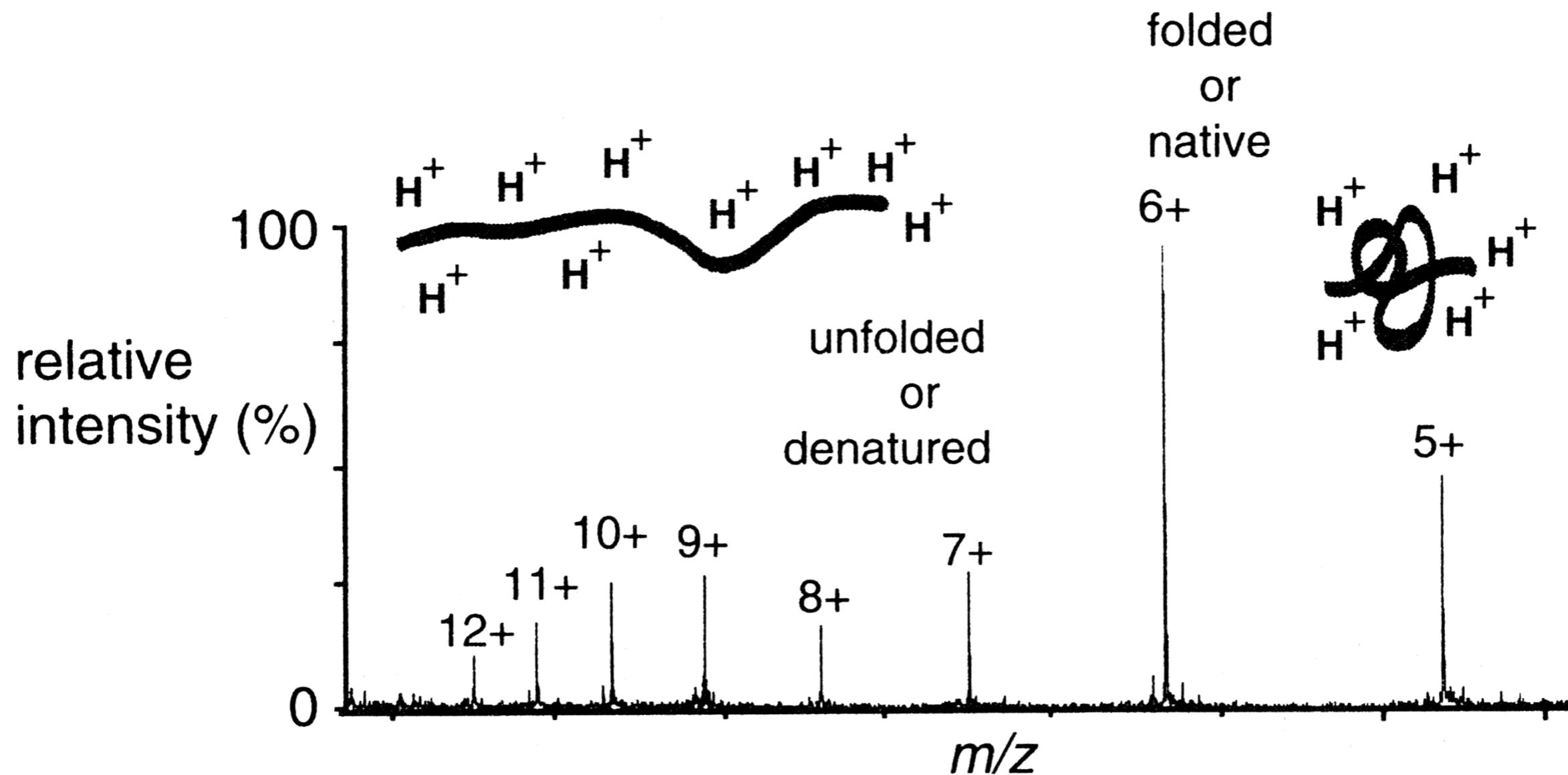
- Electrospray mass spectra show multiple ‘charge states’ for a single protein

# Determining a protein mass from an ESI spectrum

---

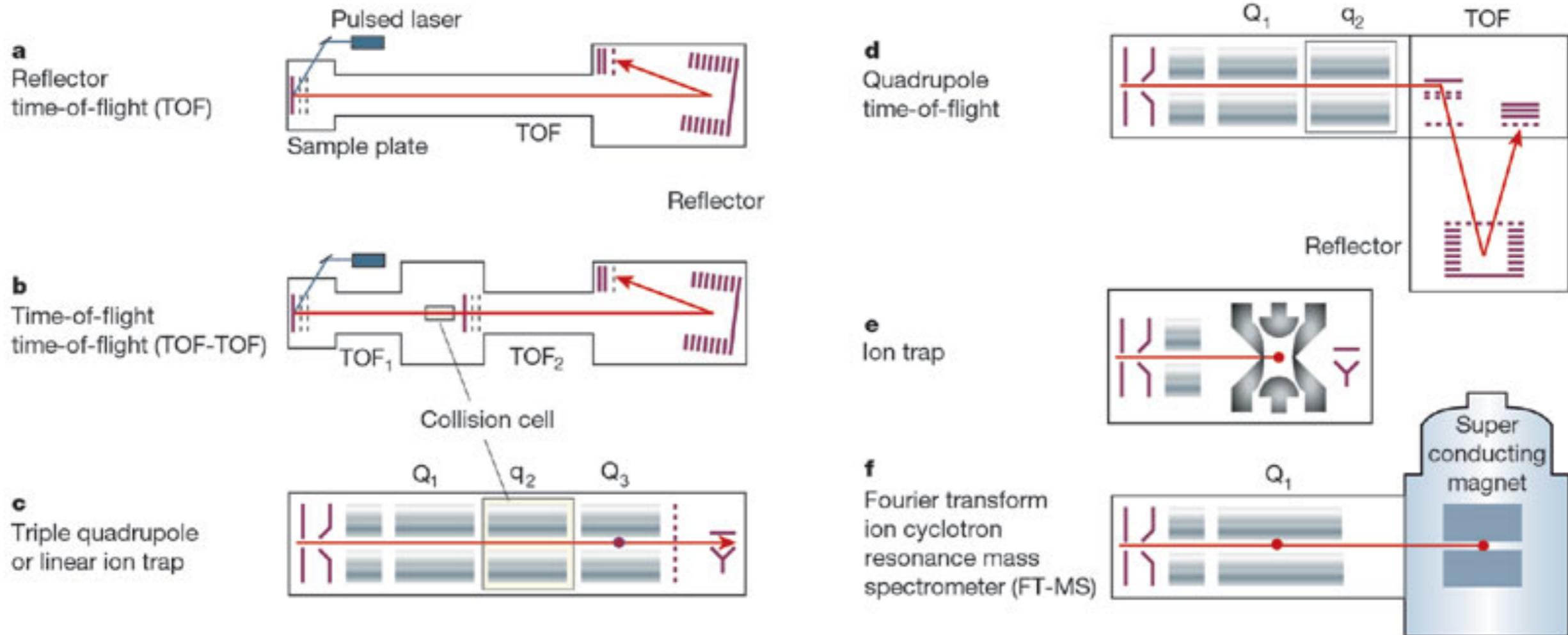
- From the  $m/z$  values of adjacent charge states, both  $m$  and  $z$  can be obtained

# Conformational effects on ESI spectra



- Folded state of protein governs its surface area, and number of sites available for protonation

# Mass spectrometer geometries

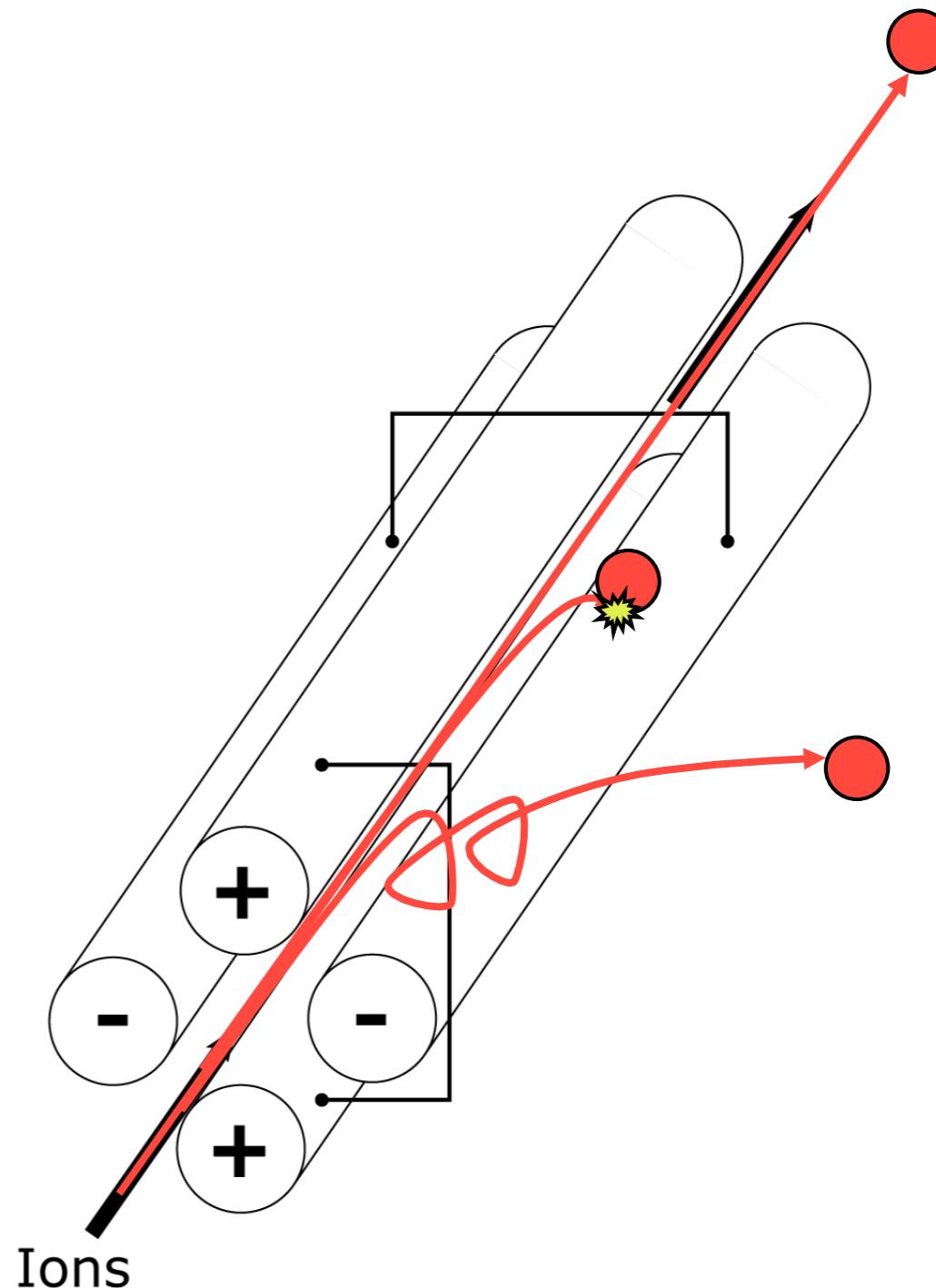


- Many mass analysis devices: quadrupole, ion trap, time-of-flight
- Mass spectrometers can contain multiple analysers in series

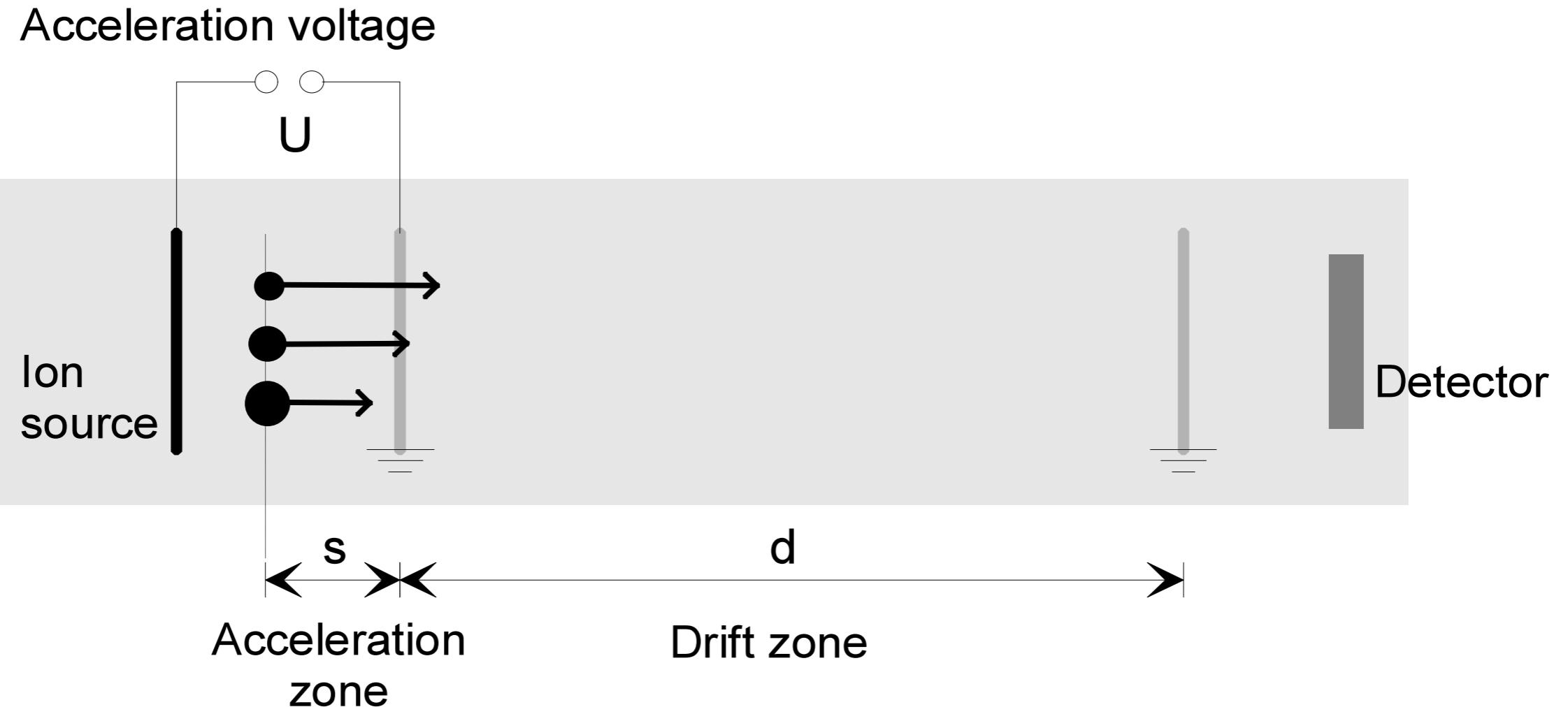
# Quadrupole (Q) analyser

---

- Four metal rods with adjacent rods at opposite polarity
- Direct current (amplitude U) and radio frequency (amplitude V) components
- Specific values of U and V allow a specific ion to traverse the quadrupole, so can be used as a filter
- Scan U and V to see ions in succession to generate complete mass spectrum



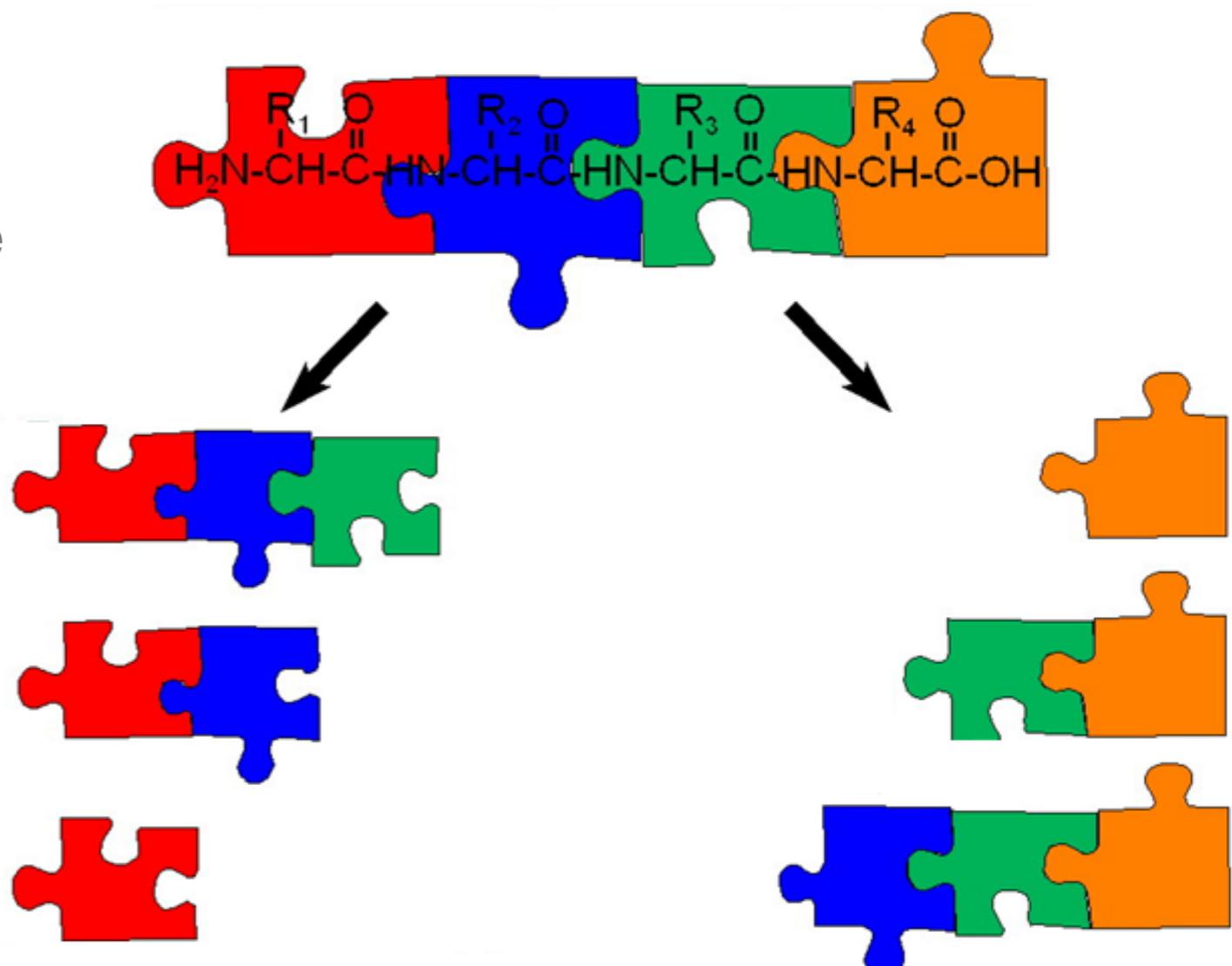
# Time-of-flight (ToF) analyser



- Ions accelerated to same kinetic energy
- Velocity, and time-of-flight, depend on  $m/z$

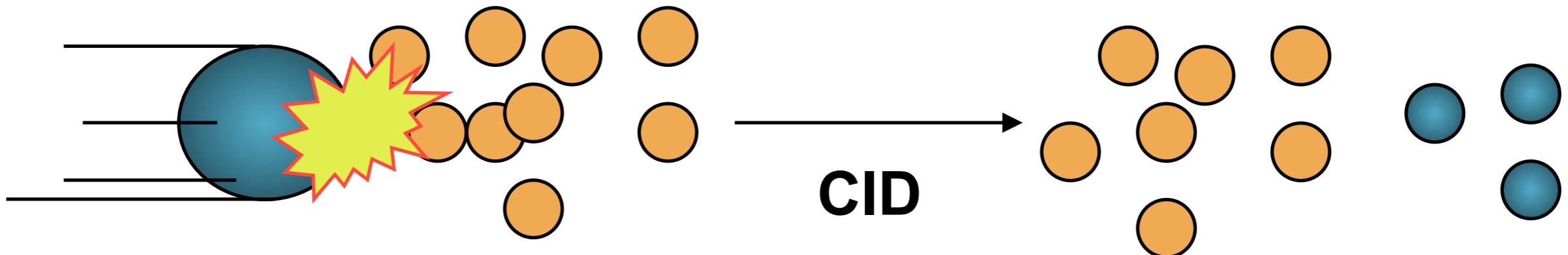
# Ion fragmentation to elucidate composition

- Intentional fragmentation of ions to assess the composition of biomolecules
- Many activation techniques have been implemented including electron-capture dissociation; infra-red multi-photon dissociation, surface-induced dissociation
- Most commonly employed is collision-induced dissociation
- Carried out in specific region of mass spectrometer



# Collision-induced dissociation (CID)

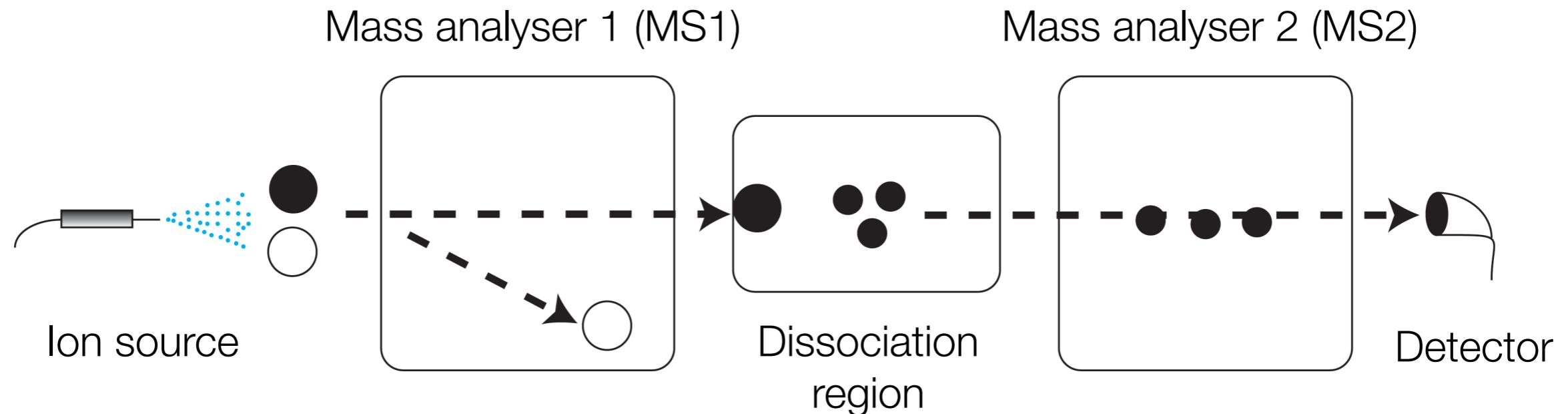
---



- Ions are accelerated into a cell containing neutral and inert gas (e.g. Ar)
- Upon collision some of the ions' kinetic energy is converted into internal modes
- The excited ion can decay into fragments, detectable on timescale of measurement

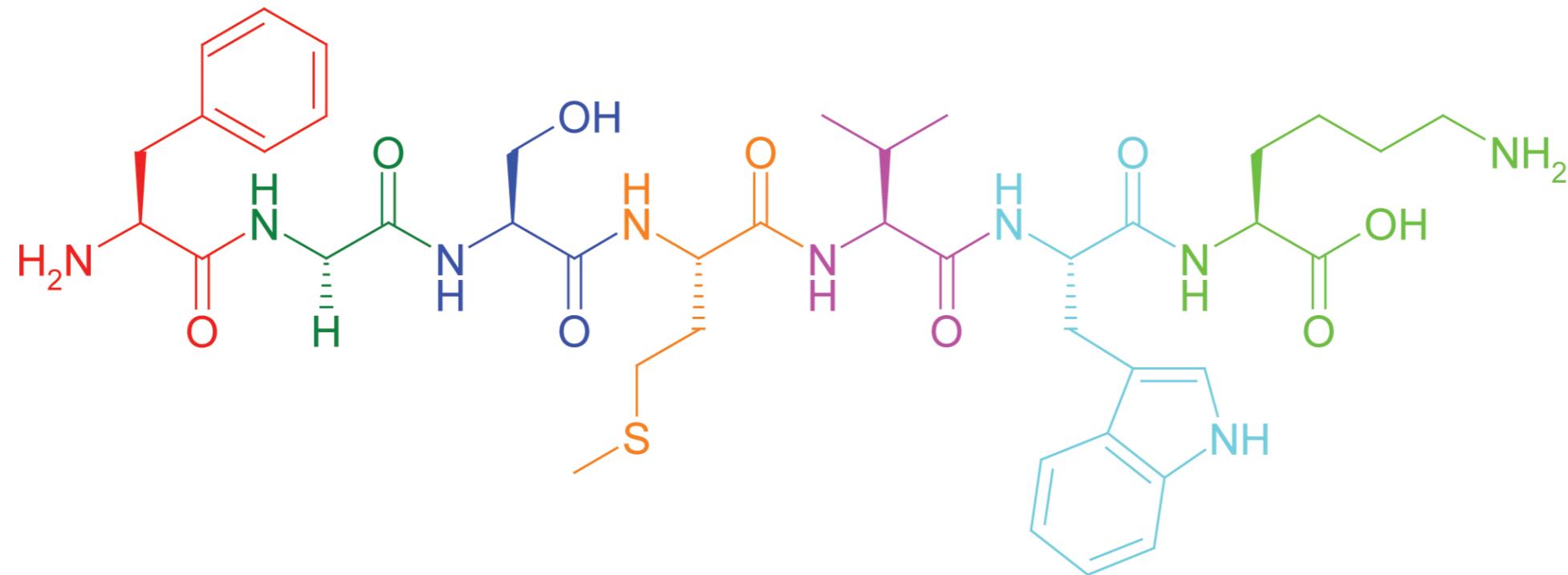
# Tandem mass spectrometry (MS/MS)

---



- The first analyser is used to select an ion population (precursor/parent ions)
- The second analyser separates dissociation products (product/daughter ions)
- The spectra obtained are often referred to as MS/MS or  $MS^2$  spectra
- Many different instrument geometries to achieve this: e.g. Q-ToF, ToF-ToF

# Peptide and protein sequence

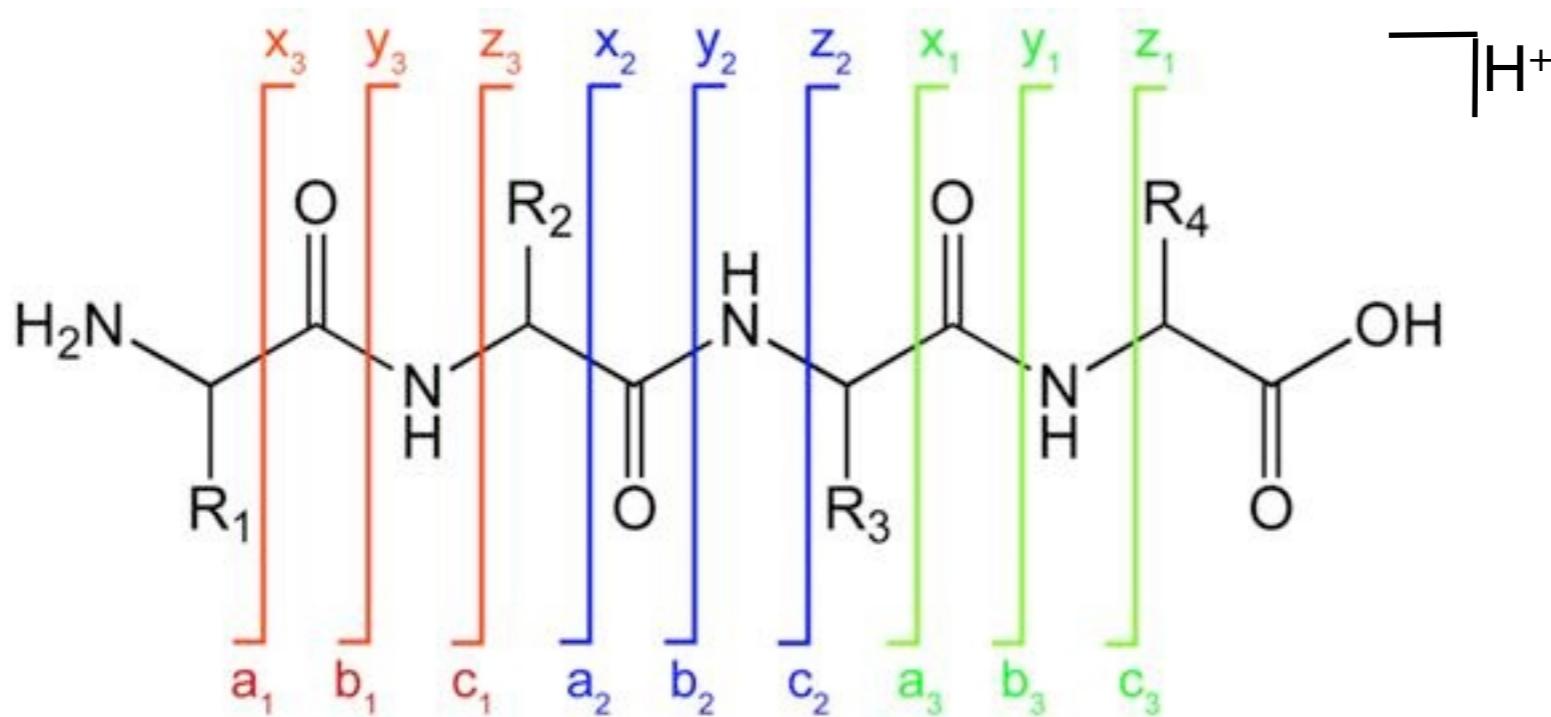


Residue 1 Residue 2 Residue 3 Residue 4 Residue 5 Residue 6 Residue 7  
Phe Gly Ser Met Val Trp Arg

- The amino acid sequence of a protein is fixed, and encodes its 3D structure
  - The sequence is labelled from the amino (N) to carboxy (C) termini

# Peptide sequencing

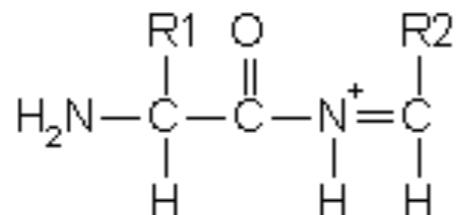
---



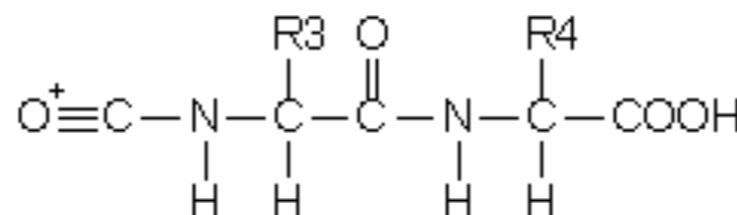
- Peptide backbone can cleave at three bonds giving different possible pairs of fragments: a and x; b and y; or c and z ions.
- a, b, c ions contain the N-terminus; x, y, z ions contain the C-terminus
- Subscript denotes the number of residues from the terminus fragmentation occurred
- Collision-induced dissociation typically results in b and y ions

# Peptide fragment ions

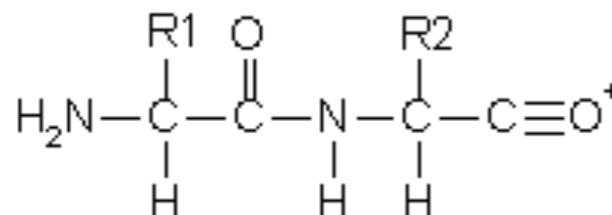
---



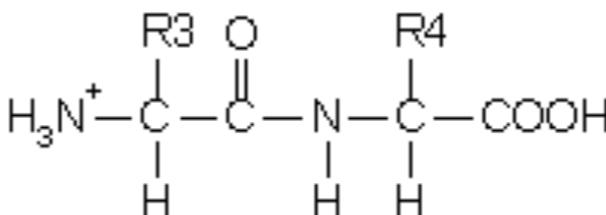
**a<sub>2</sub>**



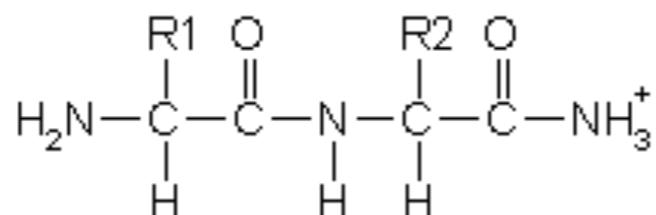
**x<sub>2</sub>**



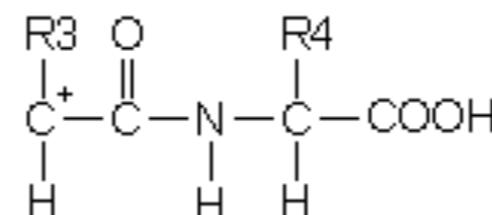
**b<sub>2</sub>**



**y<sub>2</sub>**



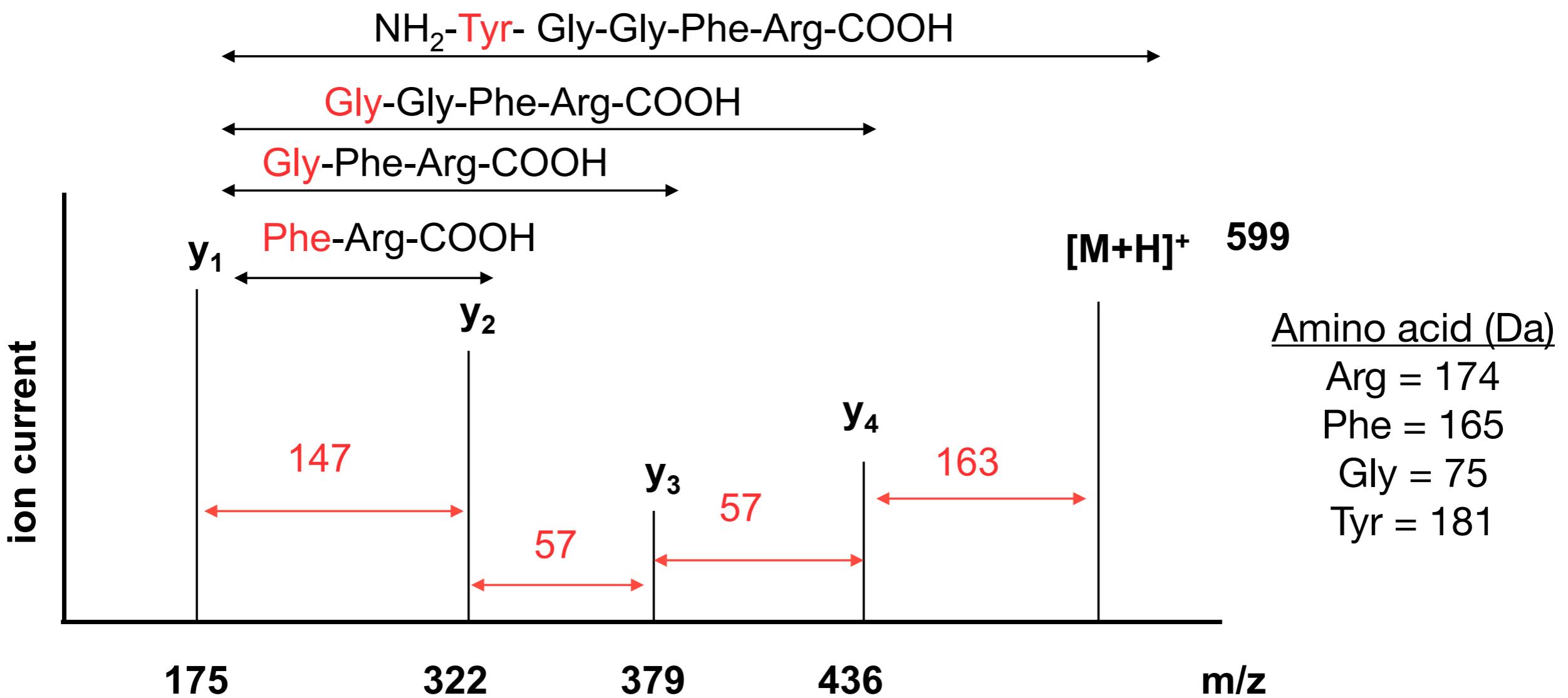
**c<sub>2</sub>**



**z<sub>2</sub>**

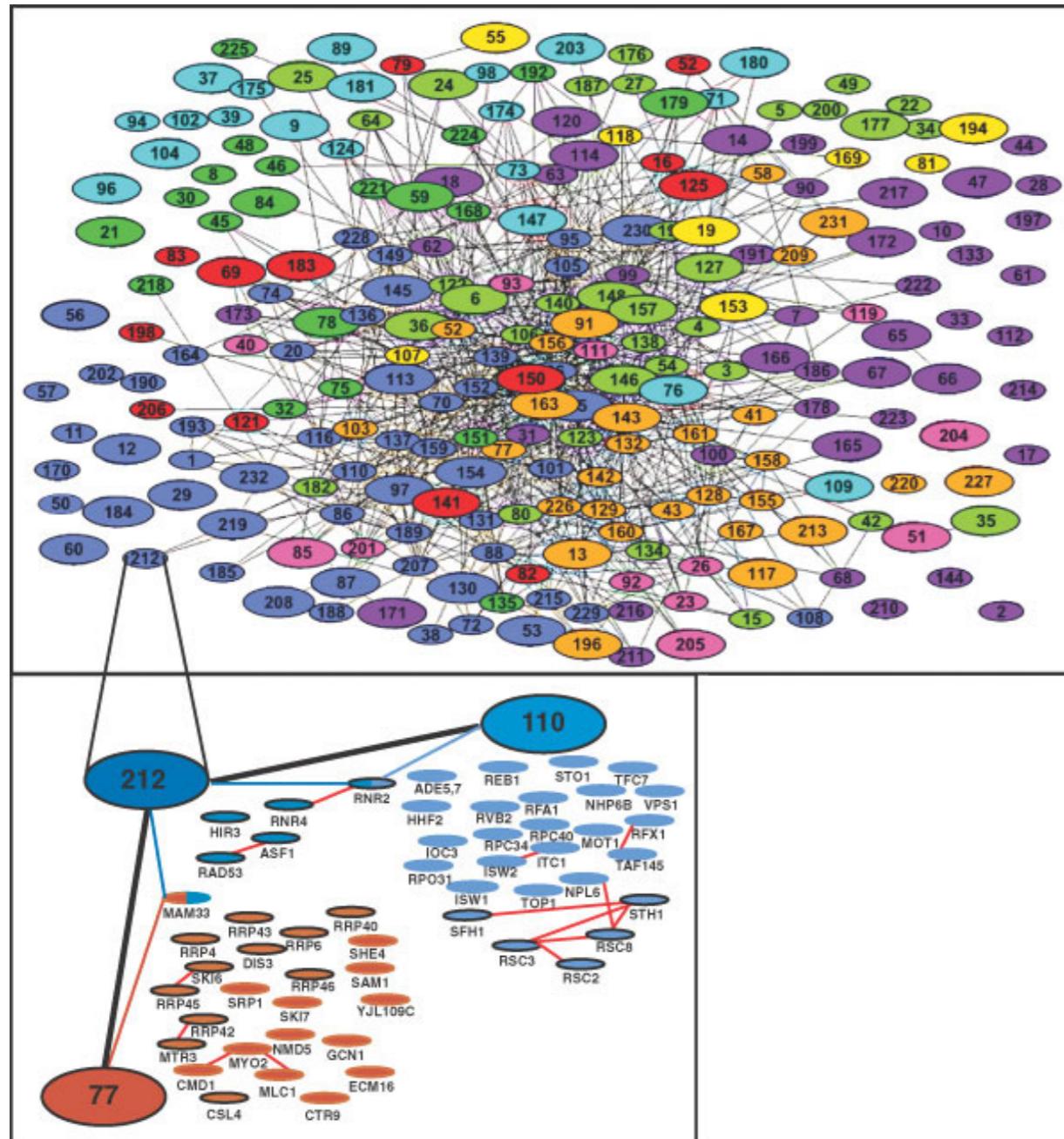
- The c and y ions abstract an additional proton from the precursor
- Multiply charged fragment ions are also often observed

# Tandem mass spectrum



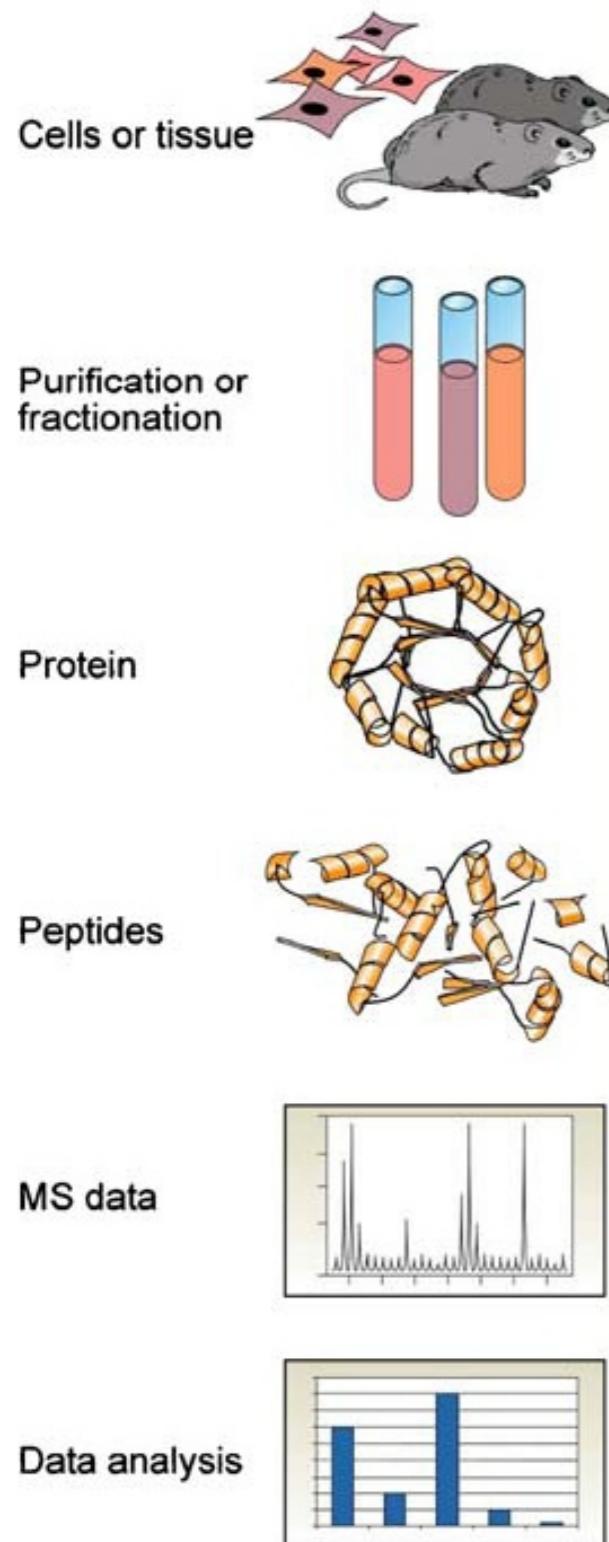
- $y_1$  is the amino acid plus 1 Da;  $b_1$  the amino acid minus 17 Da
- Sequence can be read by comparing to amino acid masses minus 18Da

# Proteomics

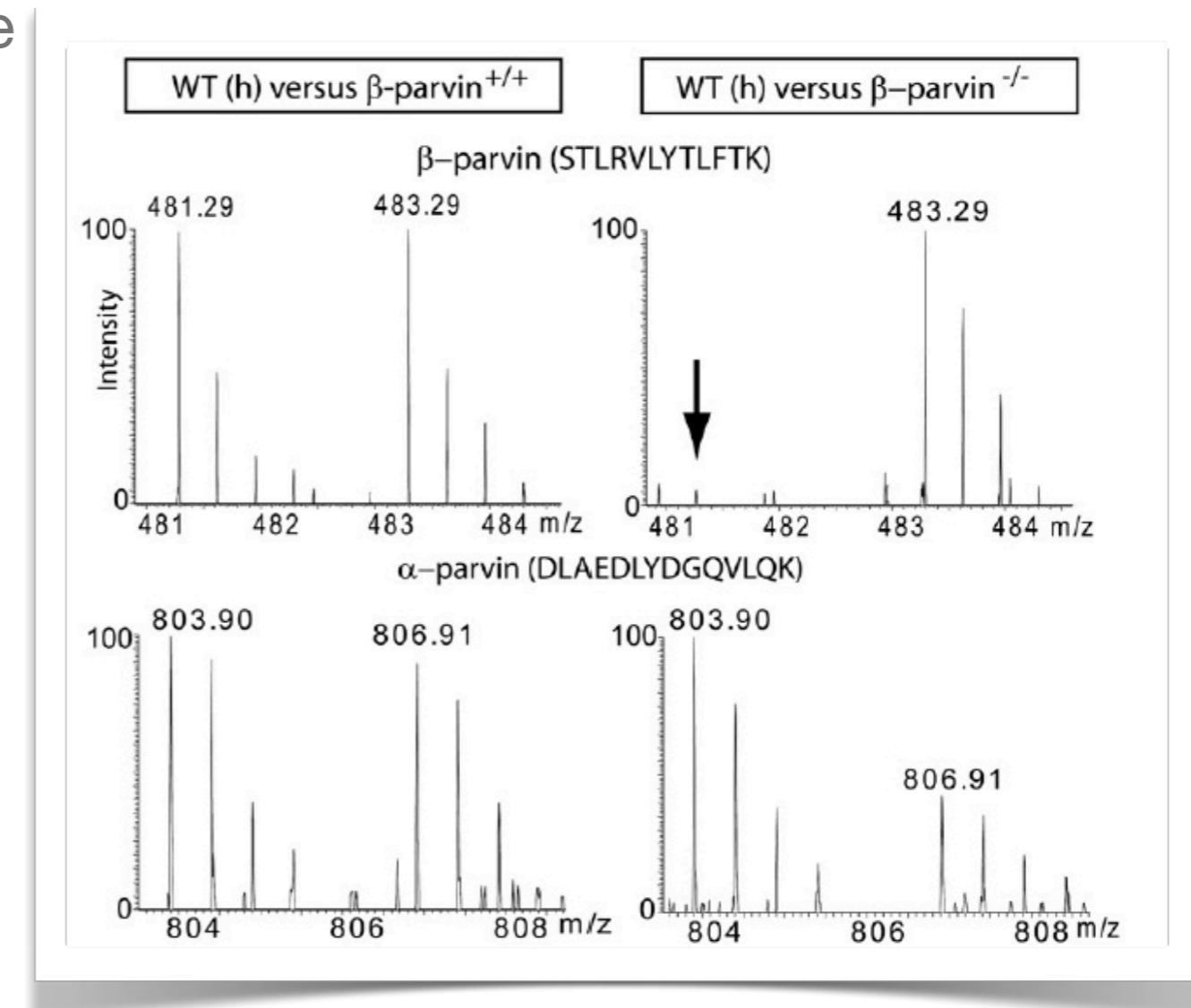


- The identification of the proteins encoded by the genome
- The determination of post-translationally modified forms
- The study of the interactions they make
- The quantification of their abundances different cells and states
- The assessment of the temporal changes they undergo

# Mass-spectrometry based proteomics - Example 1



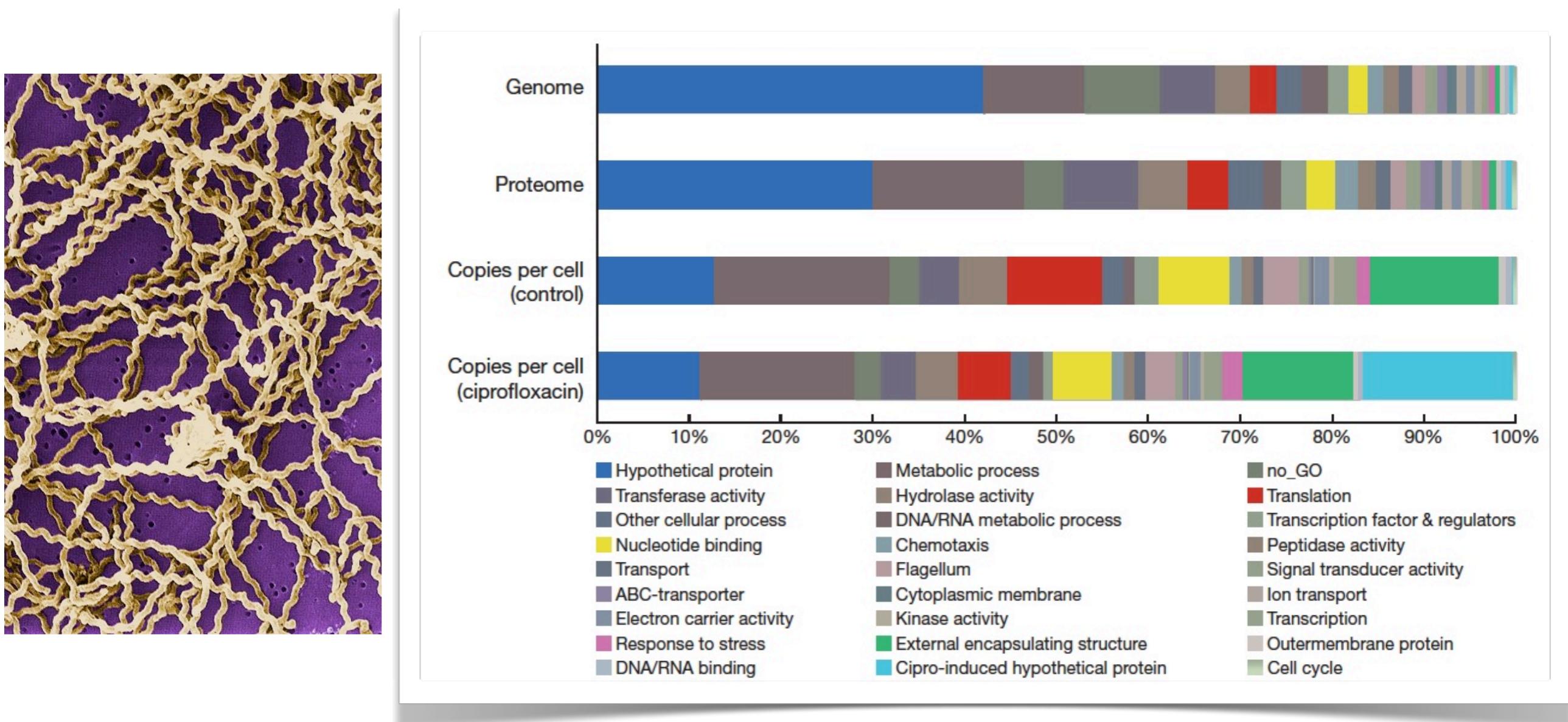
- $^{13}\text{C}$ -Lys labelled ‘normal’ mouse compared with unlabelled ‘knock-out’ mouse
- Can be used to assess proteome changes associated with disease



Kruger ... Mann, Cell (2008) 134, 353-64

# Mass-spectrometry based proteomics - Example 2

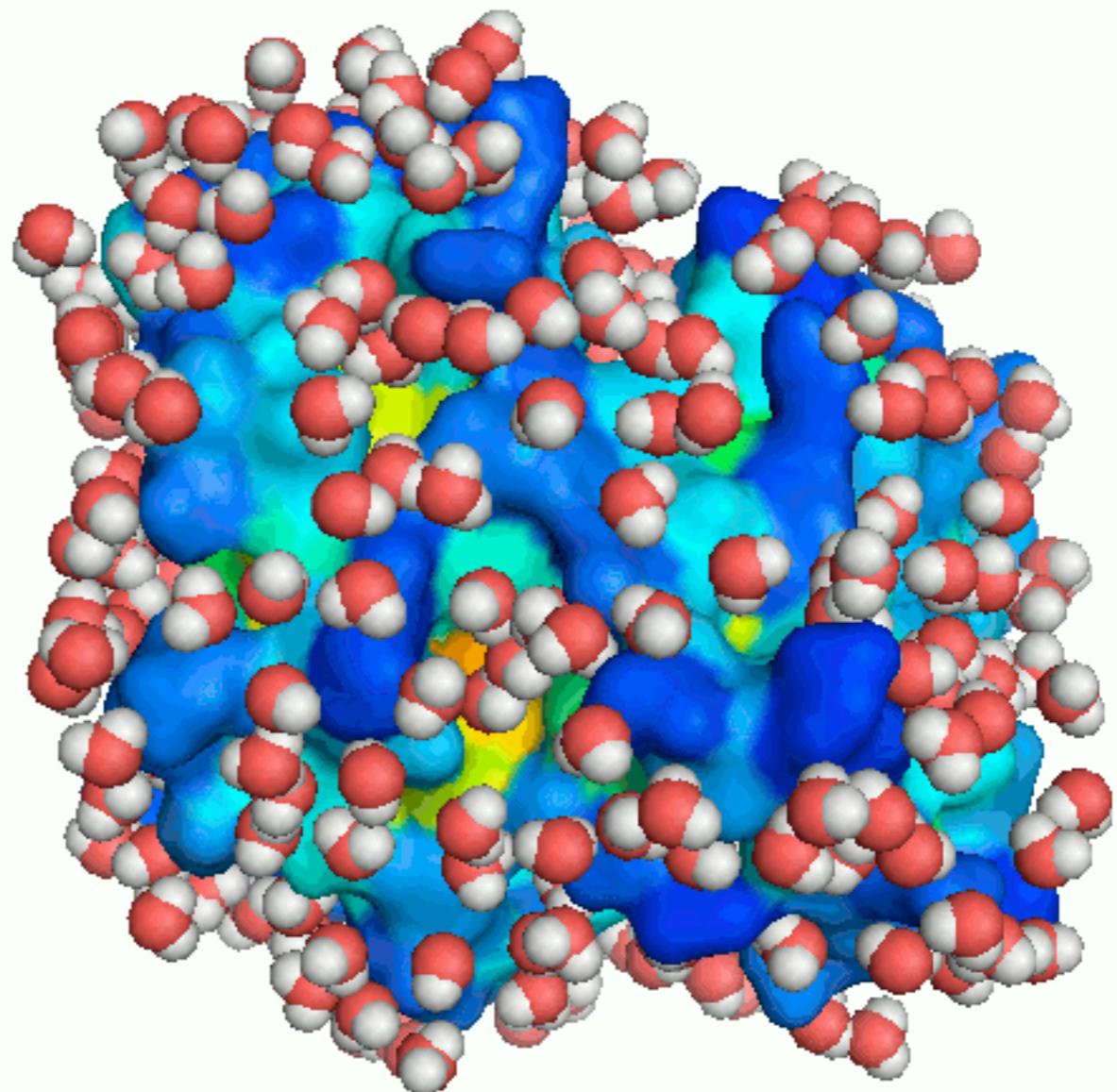
- *Leptospira interrogans* : pathogenic bacterium that causes leptospirosis
- 83% of the proteome quantified, including proteins <10 copies per cell



# Labelling technologies with MS

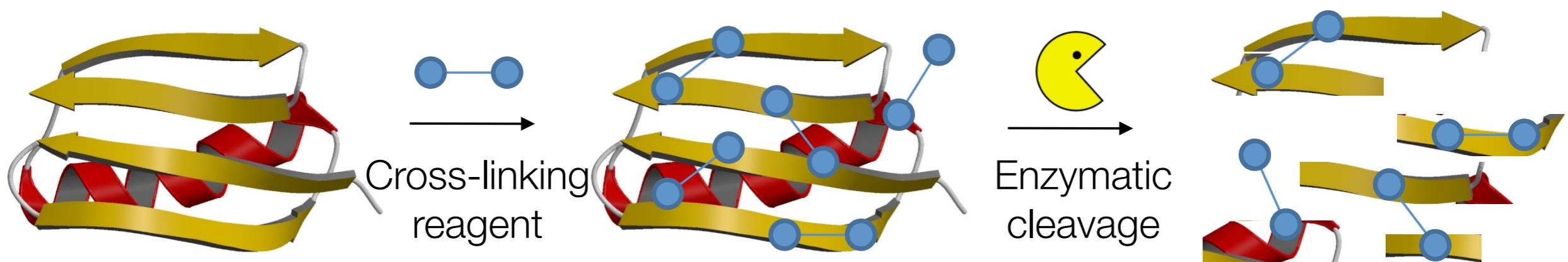
---

- Label protein, and use MS to localise the individual labels
- Can interrogate the ‘global’ and ‘local’ levels: i.e. protein or peptide level
- Peptide level accessed by enzymolysis of protein, and further interrogation with tandem MS
- Labels reveal solvent accessibility and connectivity
- Provide means of probing protein structure and dynamic fluctuations



# Cross-linking mass spectrometry

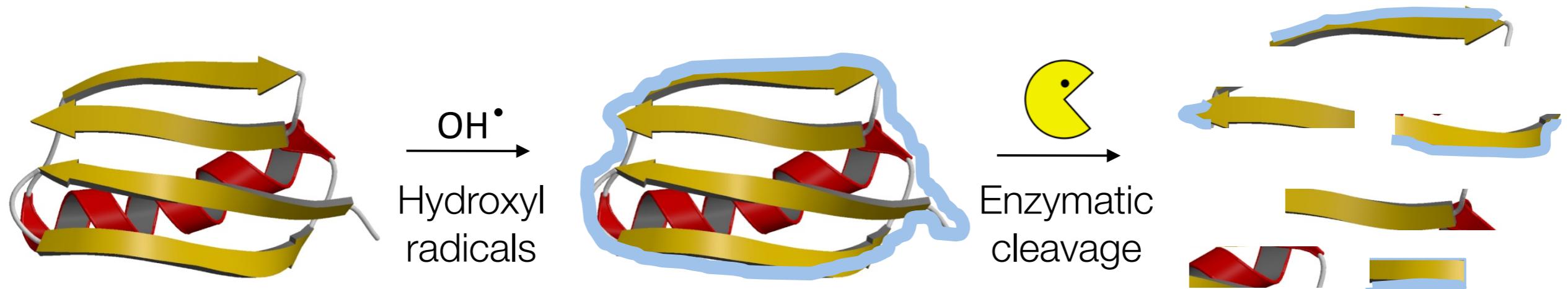
---



- Cross-linker forms covalent bonds between amino acids with appropriate functionality
- Peptide masses (and fragments) are interrogated to localise cross-linker
- Intra- and inter-protein cross-links can be formed
- Cross-links can be used to determine connectivity, and as a spatial restraint for modelling protein structures

# Oxidative foot-printing

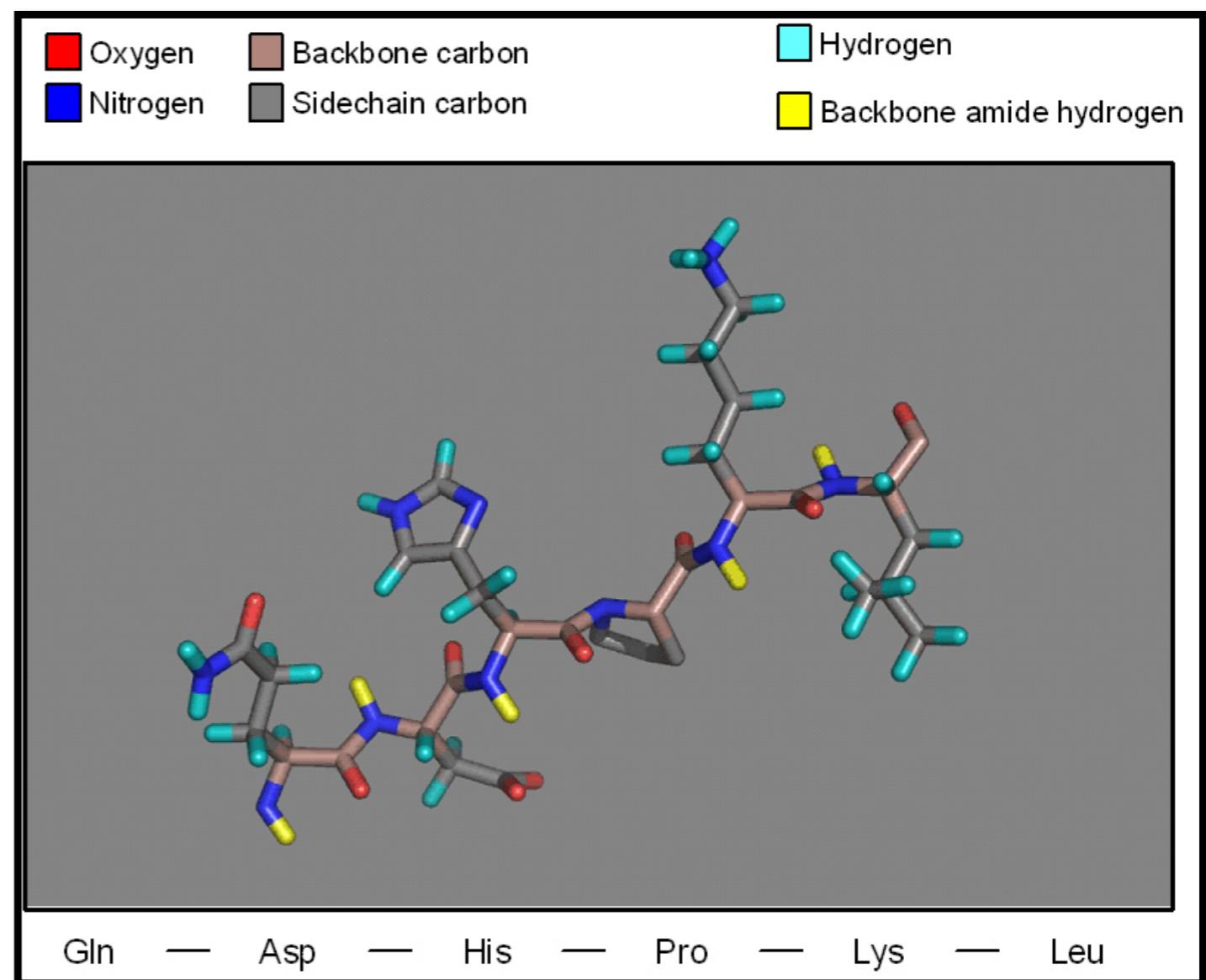
---



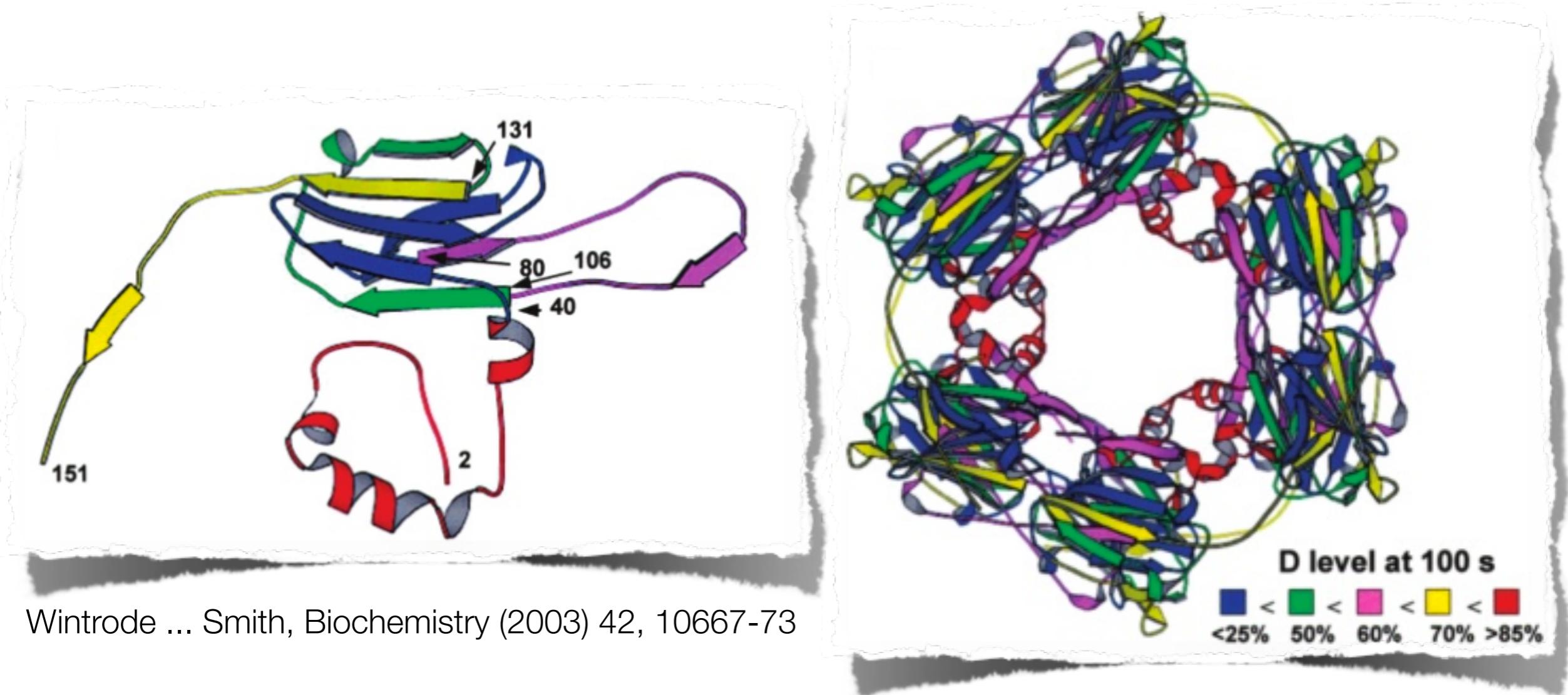
- Solvent accessible amino acid side chains of the protein are oxidised
- Peptide masses (and fragments) are interrogated to localise oxidation sites
- Comparing data from proteins in complex and in isolation allows the determination of interface sites

# Hydrogen/deuterium exchange

- Monitor the rate at which protein hydrogens are replaced by deuteriums (or vice versa)
- Three types of hydrogen in proteins, only backbone amide hydrogens exchange at measurable rate
- Exchange can be (effectively) quenched by dropping pH to ~2.5, and temperature to 0°C
- Exchange rates reveal solvent accessibility



# Labelling technology - Example



- Hydrogen/deuterium exchange of oligomeric ‘molecular chaperone’ protein
- Side-chains at interfaces exchange relatively rapidly, suggesting a labile oligomer

# Tomorrow - Gas-Phase Structural Biology

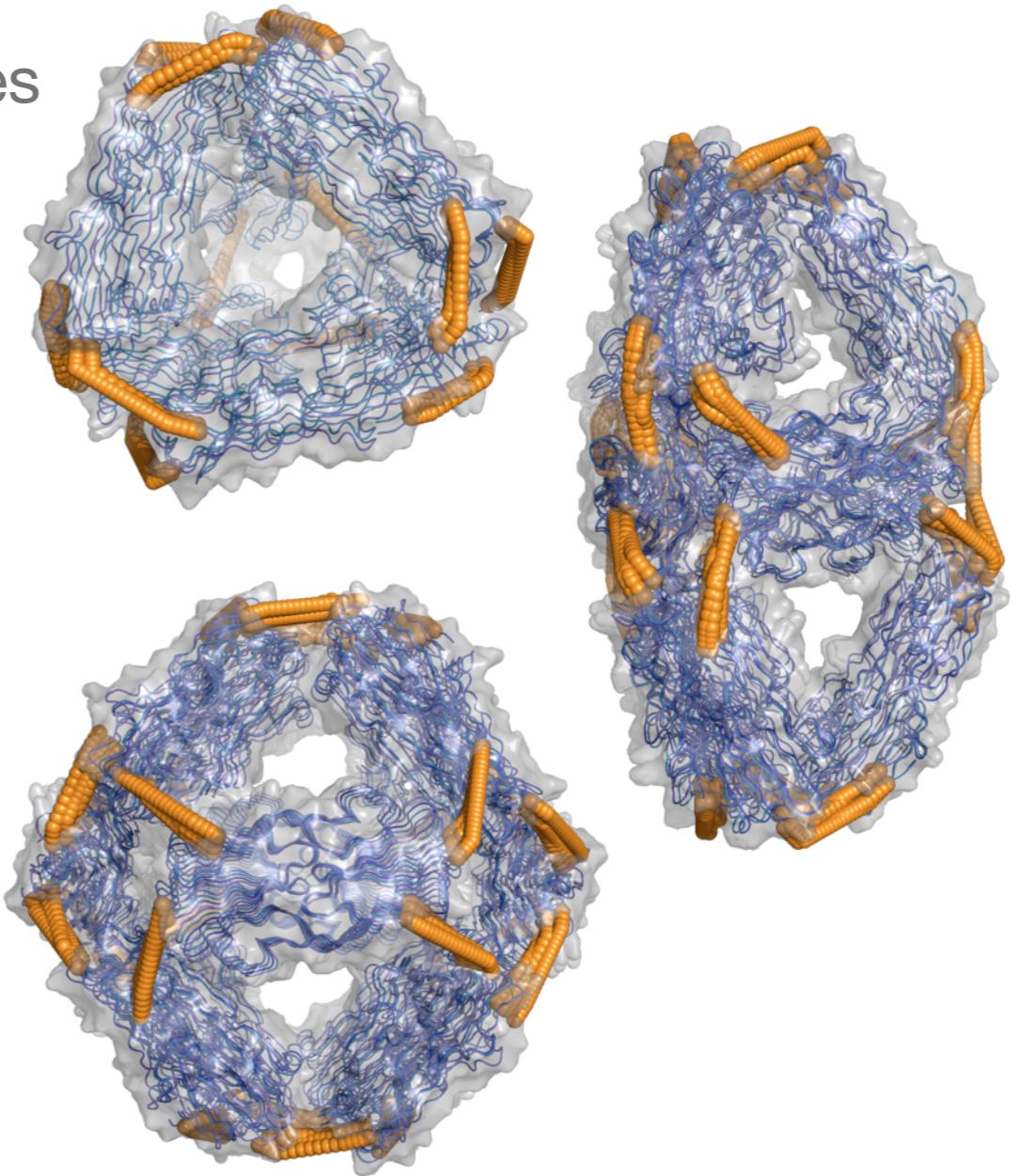
---

- **Mass spectrometry of macromolecular assemblies**

- Mass measurement of very large molecules
- Ion mobility spectrometry
- Preservation of structure in the gas phase

- **Determination of structure and dynamics**

- Protein polydispersity and heterogeneity
- Quaternary dynamics
- Hybrid approaches for structural biology



# Biophysical Chemistry

---

Justin Benesch

Jonathan Doye

Mark Wallace

# Lecture 8 - Gas-Phase Structural Biology

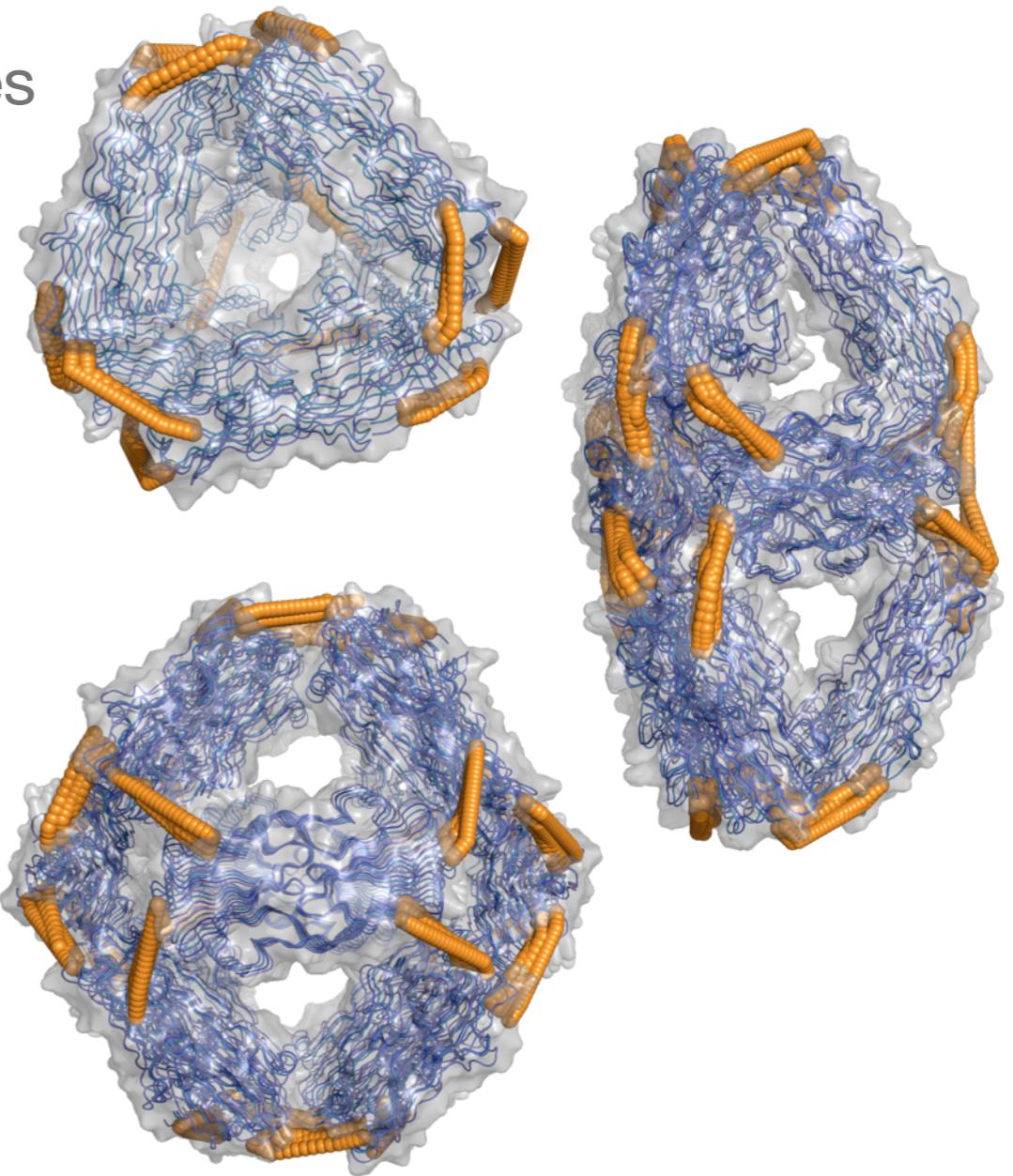
---

- **Mass spectrometry of macromolecular assemblies**

- Mass measurement of very large molecules
- Ion mobility spectrometry
- Preservation of structure in the gas phase

- **Determination of structure and dynamics**

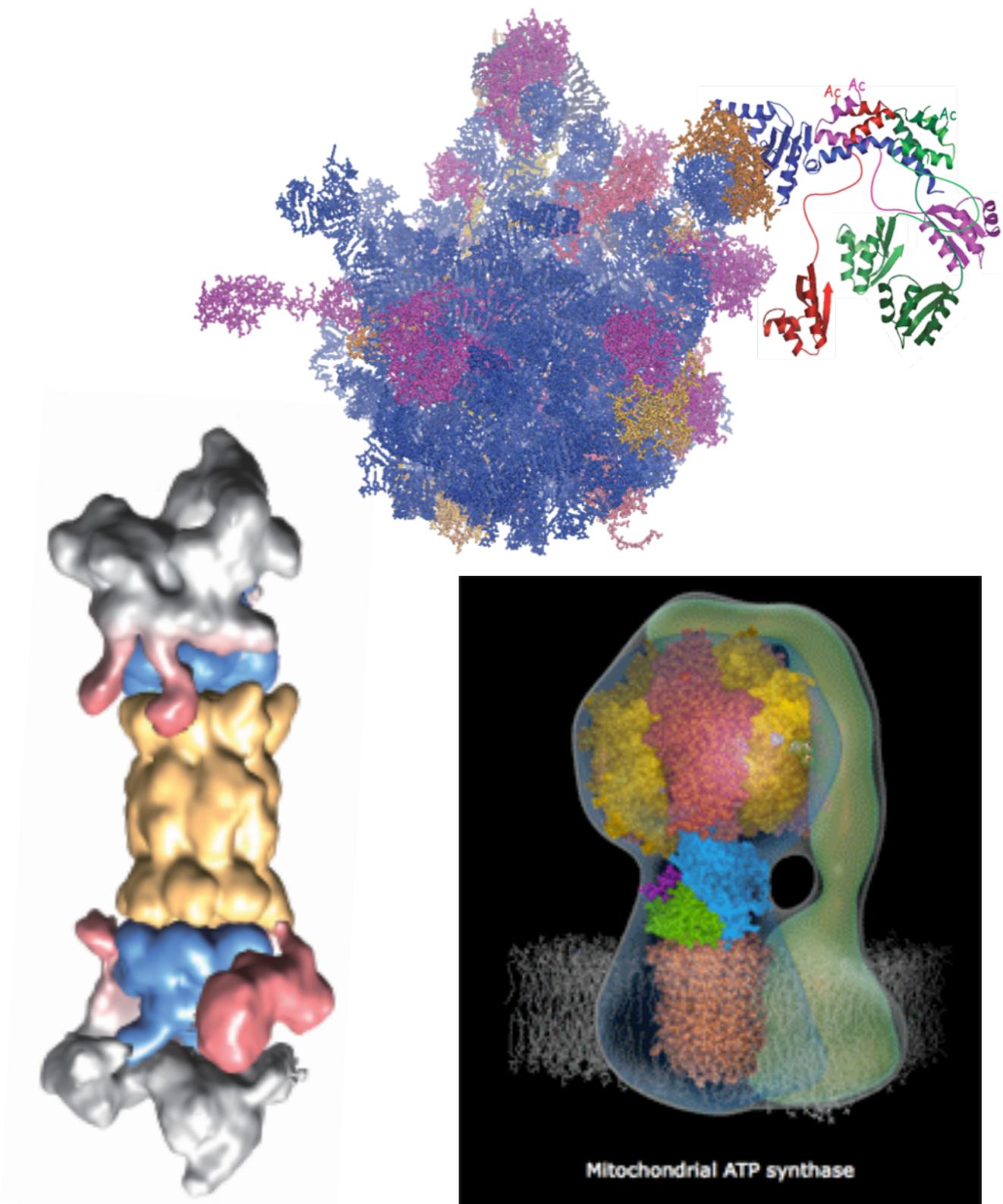
- Protein polydispersity and heterogeneity
- Quaternary dynamics
- Hybrid approaches for structural biology



# Proteins carry out their functions as assemblies

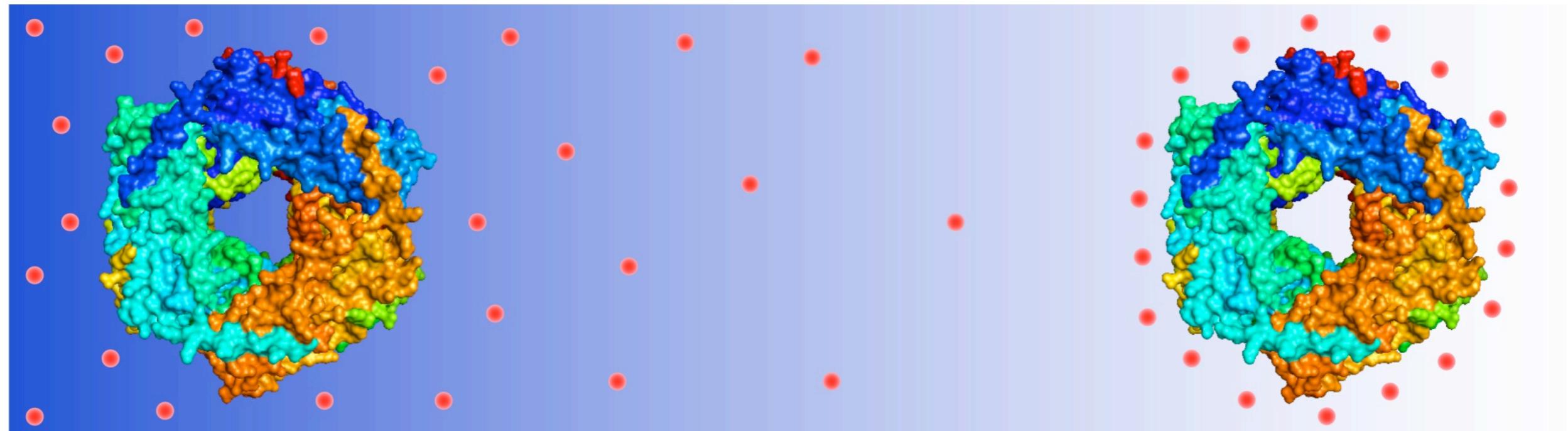
---

- Majority of proteins exist as multi-subunit complexes
- Complexes can be homomeric or heteromeric
- Average complex size is 4.9 subunits per oligomer (in yeast)
- This quaternary structure is mediated by noncovalent interactions



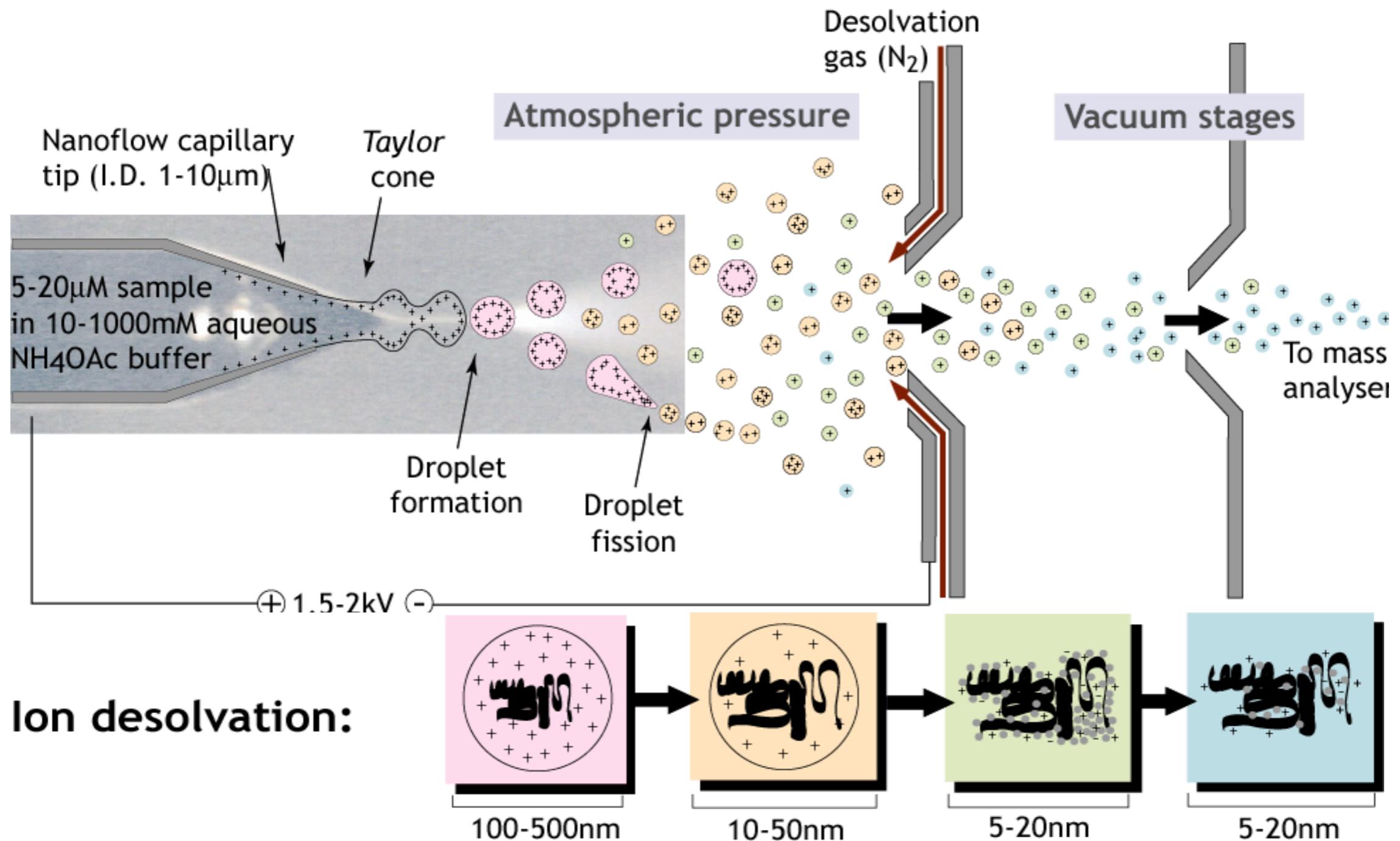
# Maintaining noncovalent interactions

---

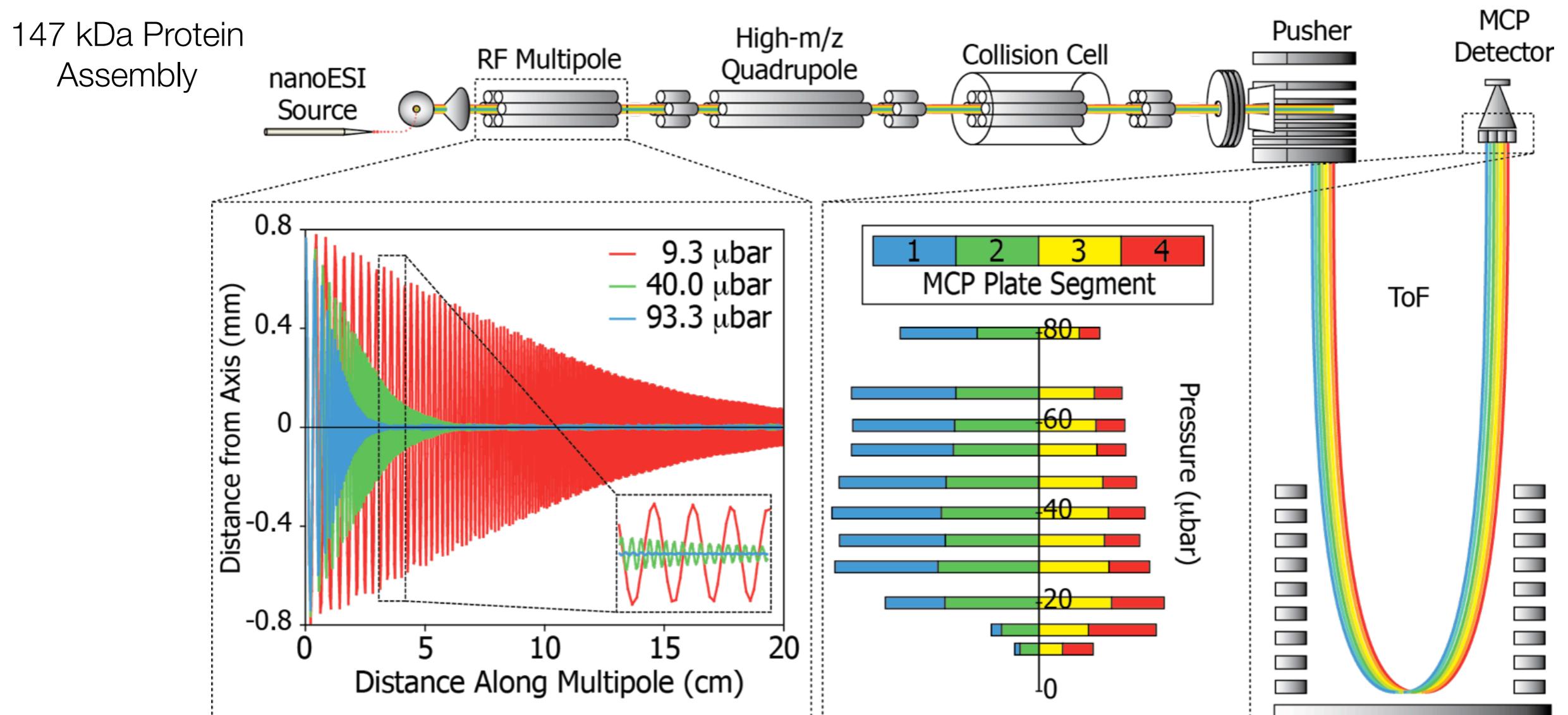


- Transfer multi-subunit protein assembly from solution into gas phase
- Requires control of ionisation conditions, and ion transmission

# Nano-electrospray ionisation (nESI)

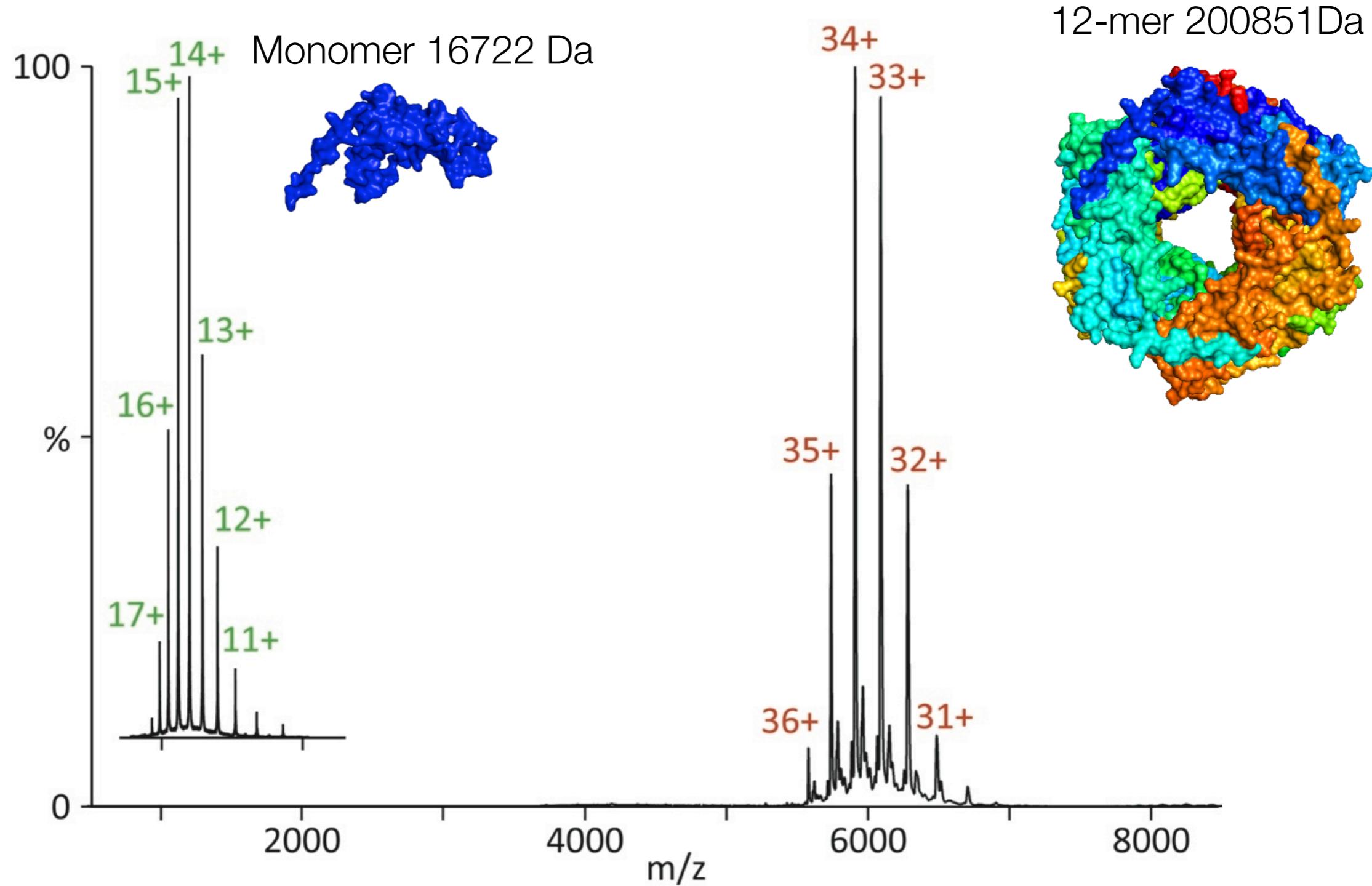


# Collisional focussing

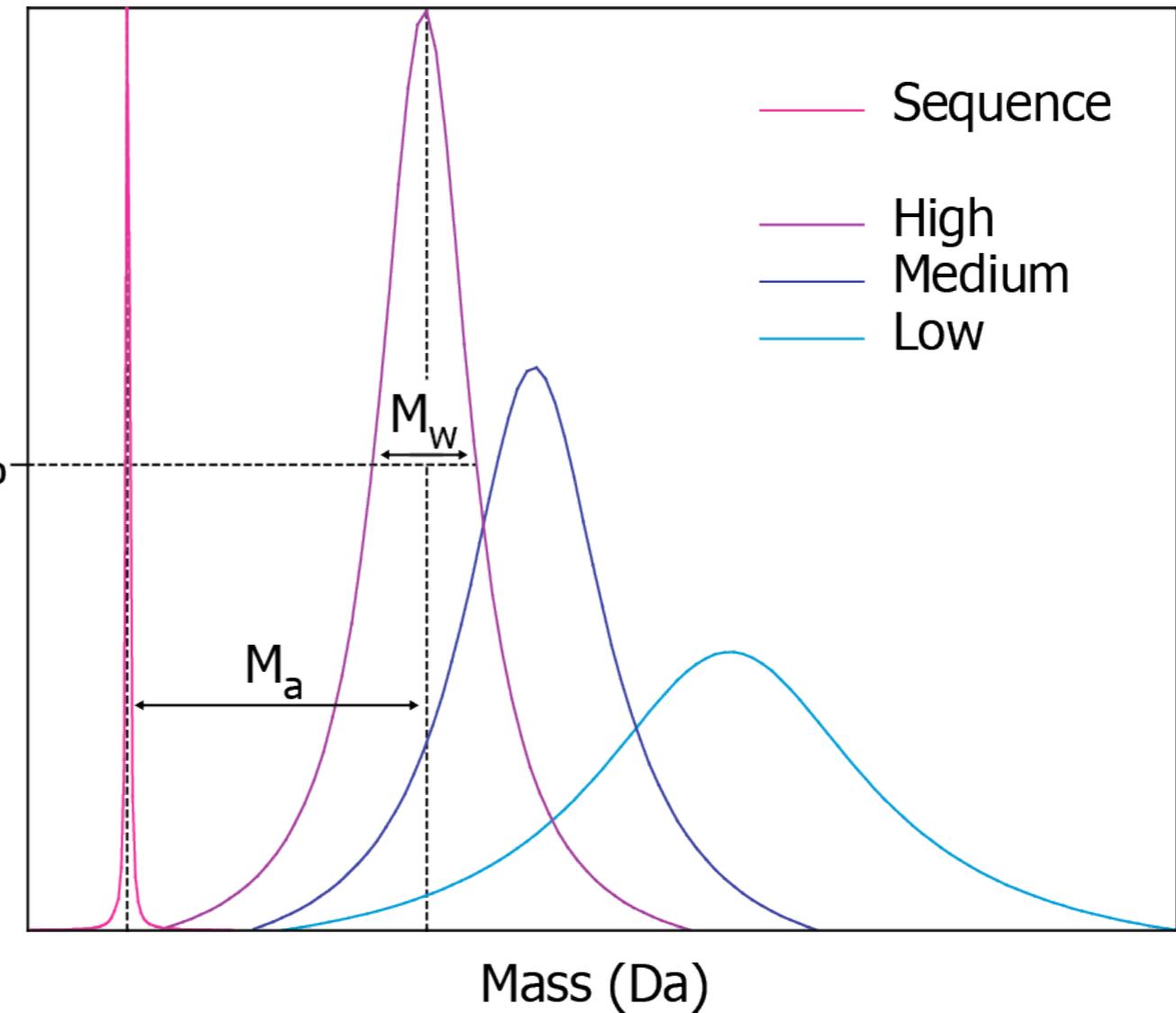
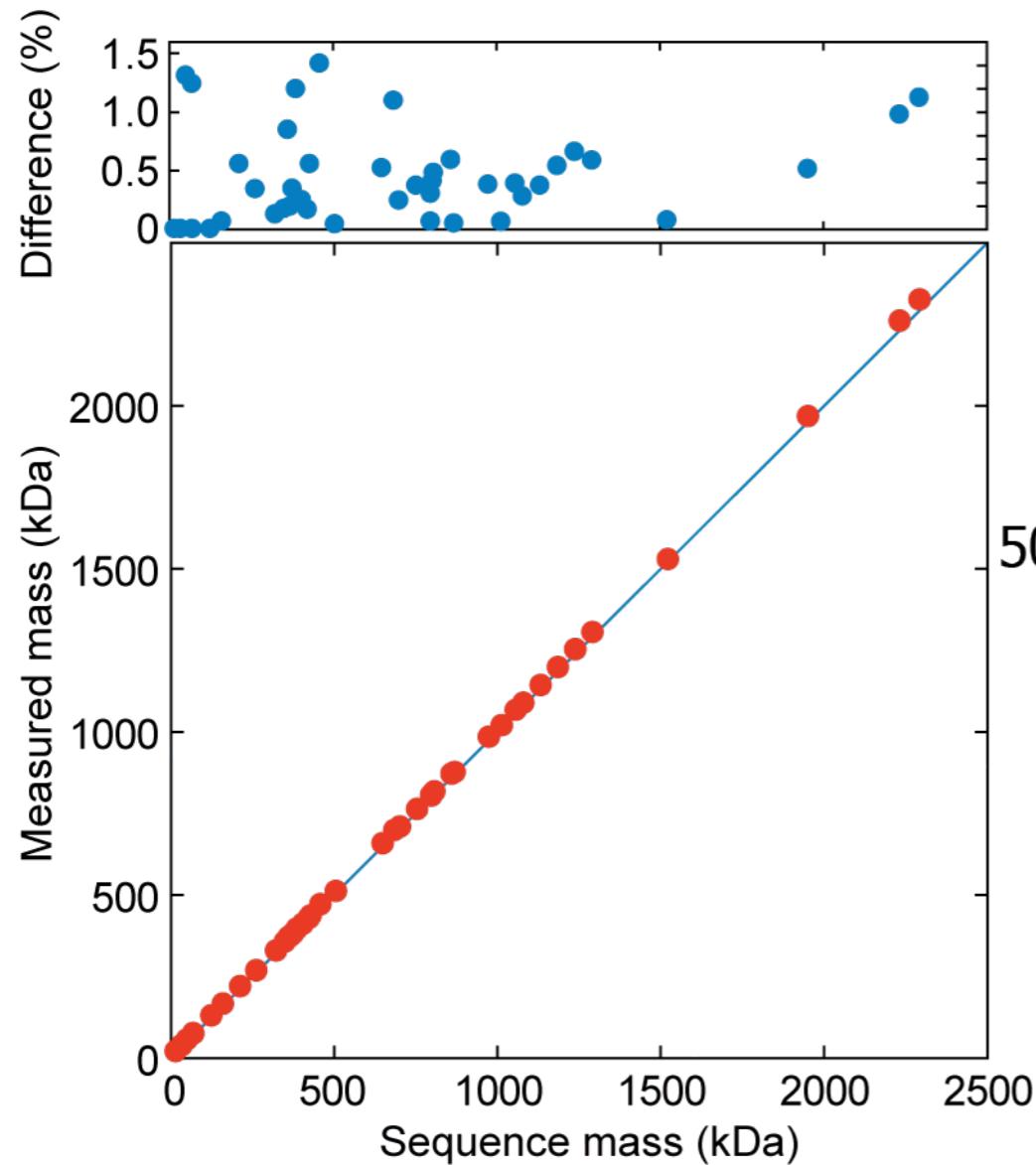


- Both axial and radial components of the ions' velocity can be damped by collisions with background gas

# nESI mass spectrum



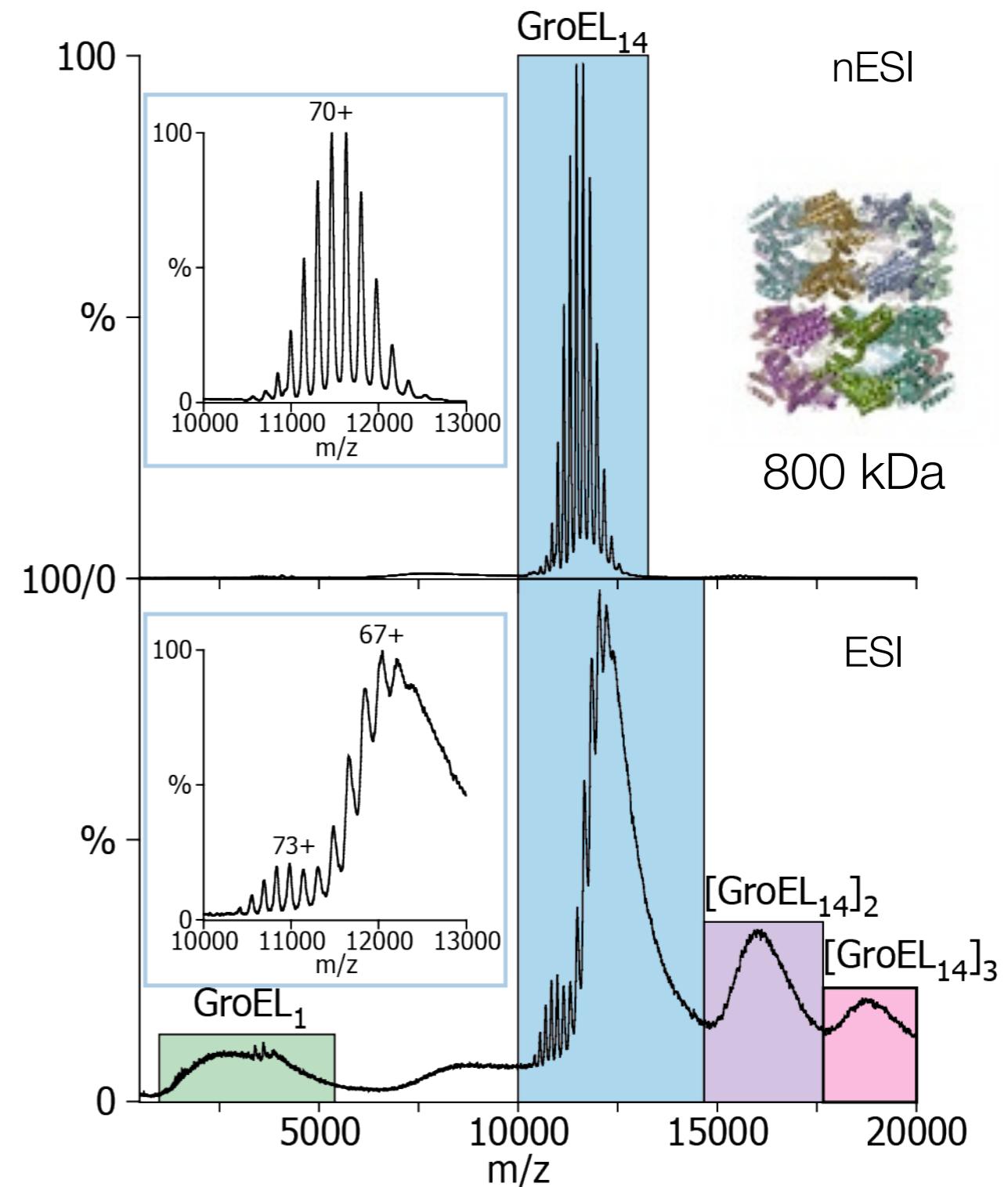
# Mass accuracy



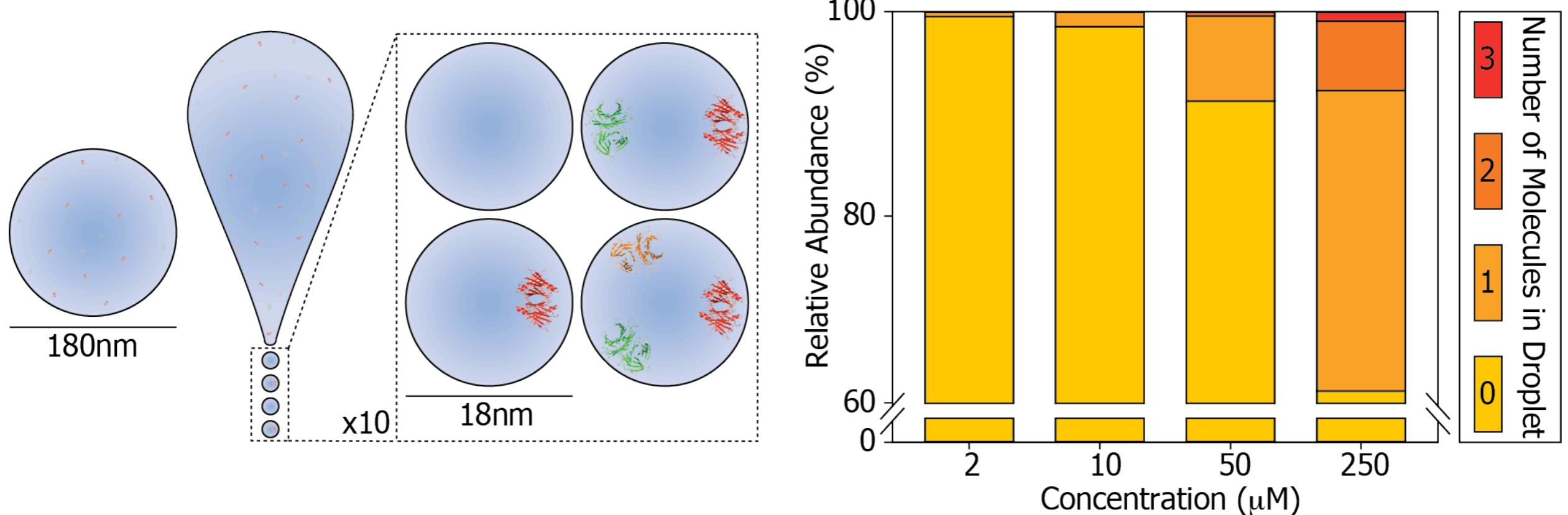
- Additional mass due to adducted solvent molecules and buffer ions
- Number of adducts inversely related to activation

# Benefits of nESI

- Lower sample amounts (flow rate approx 10nL/min, vs 5  $\mu$ L/min in ESI)
- Can use aqueous buffers and ambient temperatures
- Narrower charge states due to fewer adduction
- Less dissociation of oligomer
- Symmetrical charge state distribution indicative of a single conformation
- Fewer non-specific aggregates

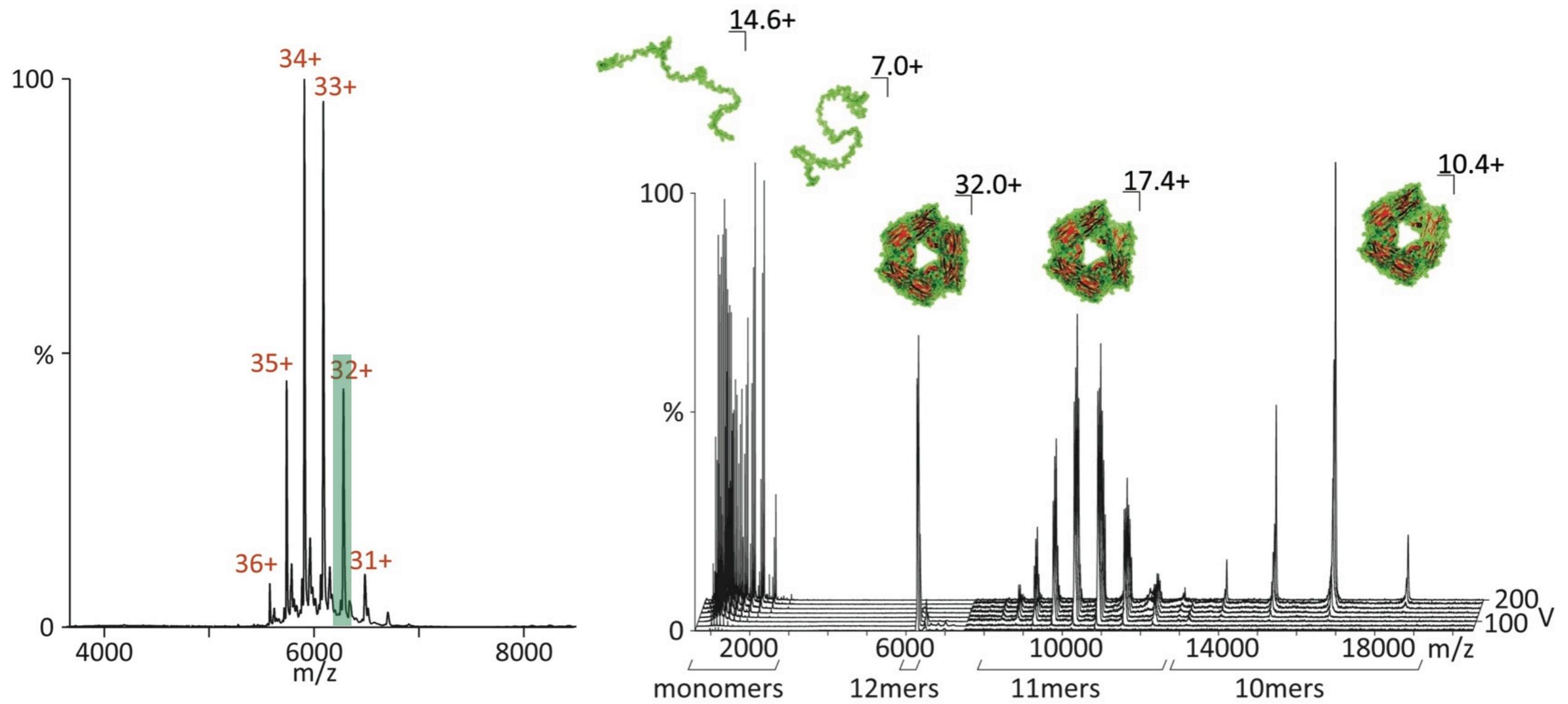


# Non-specific associations during ESI



- Probability of there being >1 analyte molecules in ‘final’ ESI droplet
- Most droplets are empty, occupancy increases with concentration
- Decreased initial droplet size in nESI reduces prevalence of non-specific aggregates

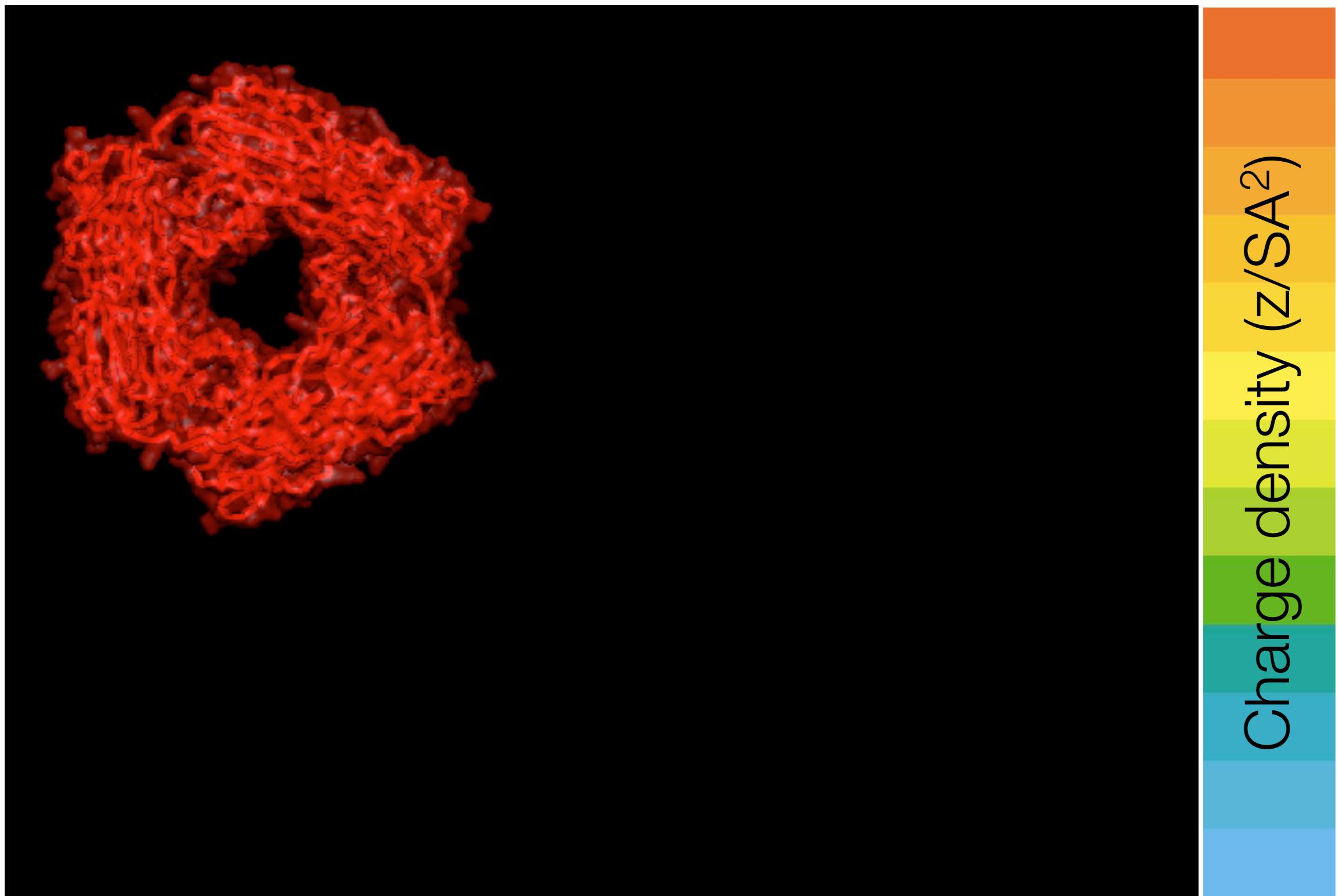
# Collision induced dissociation of protein assemblies



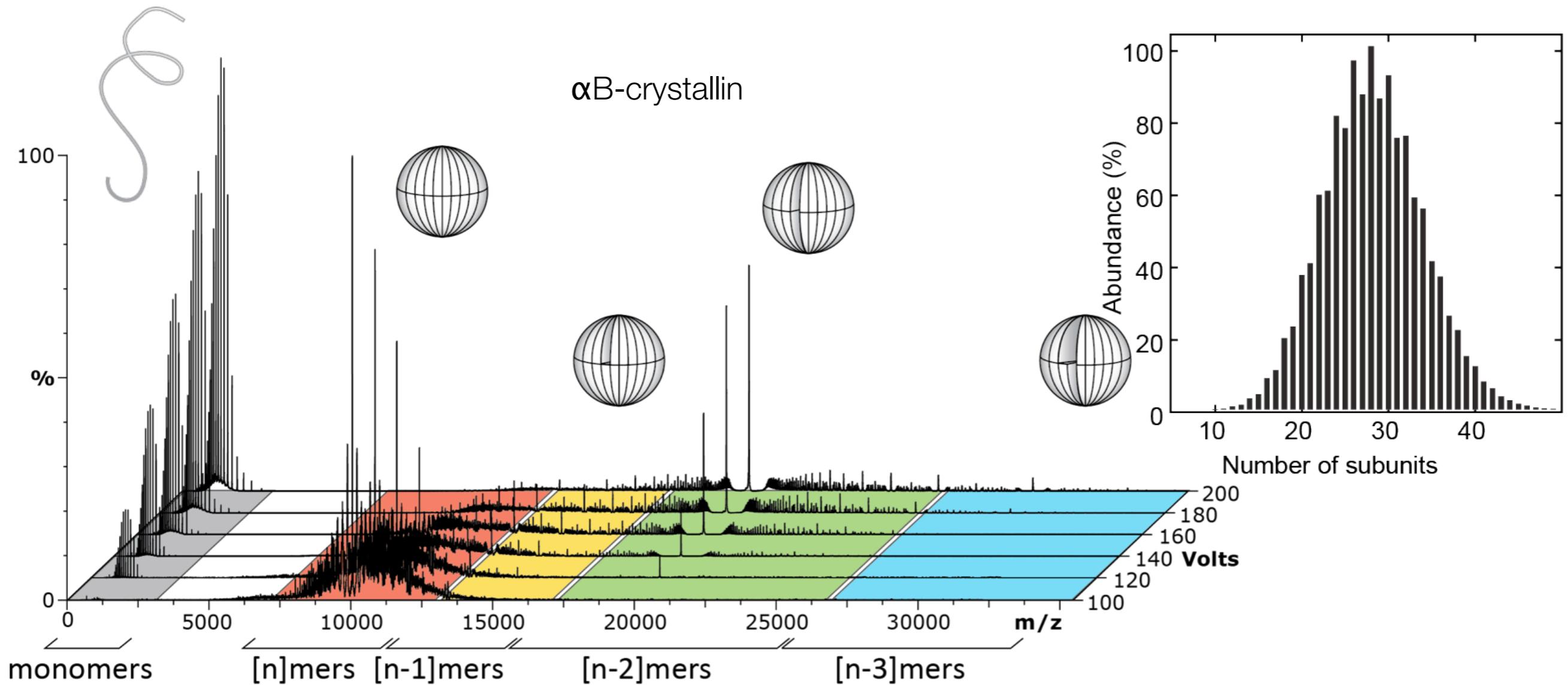
- Dissociation is asymmetric with respect to mass
- Unfolded, highly charged monomers are removed sequentially

# Dissociation causes decrease in charge density

---

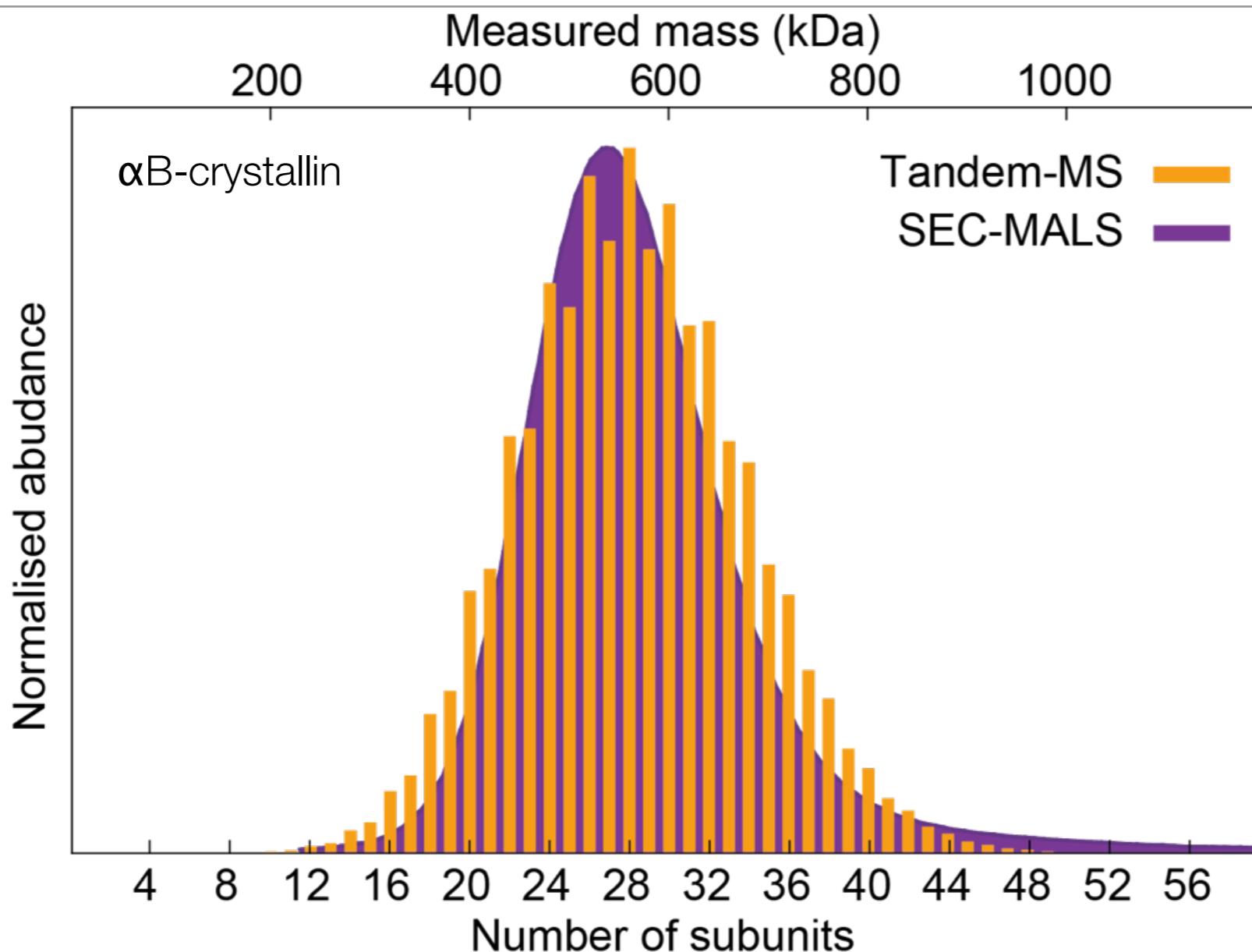


# Deconvoluting heterogeneity with CID



- Peak separation is aided by the charge reduction afforded by CID
- Predictable nature of CID allows back calculation of oligomeric distribution

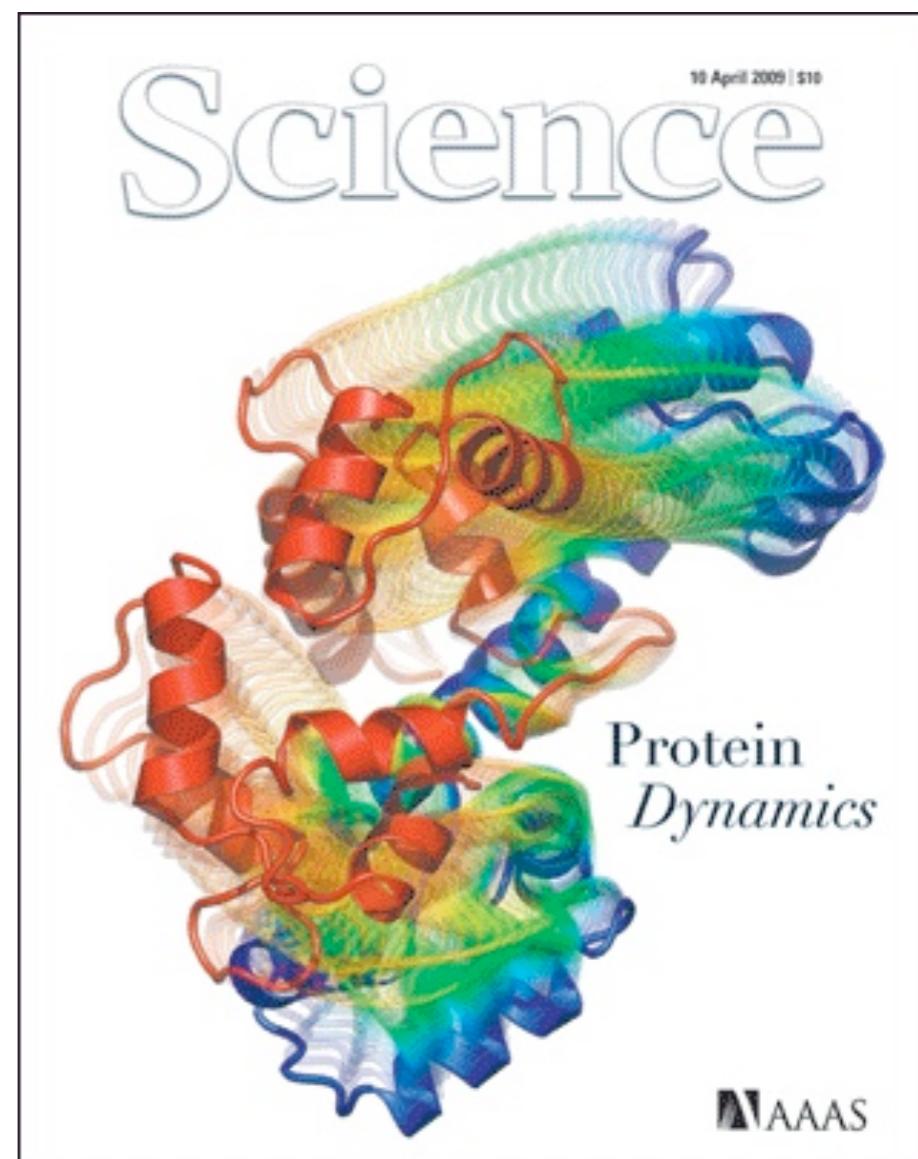
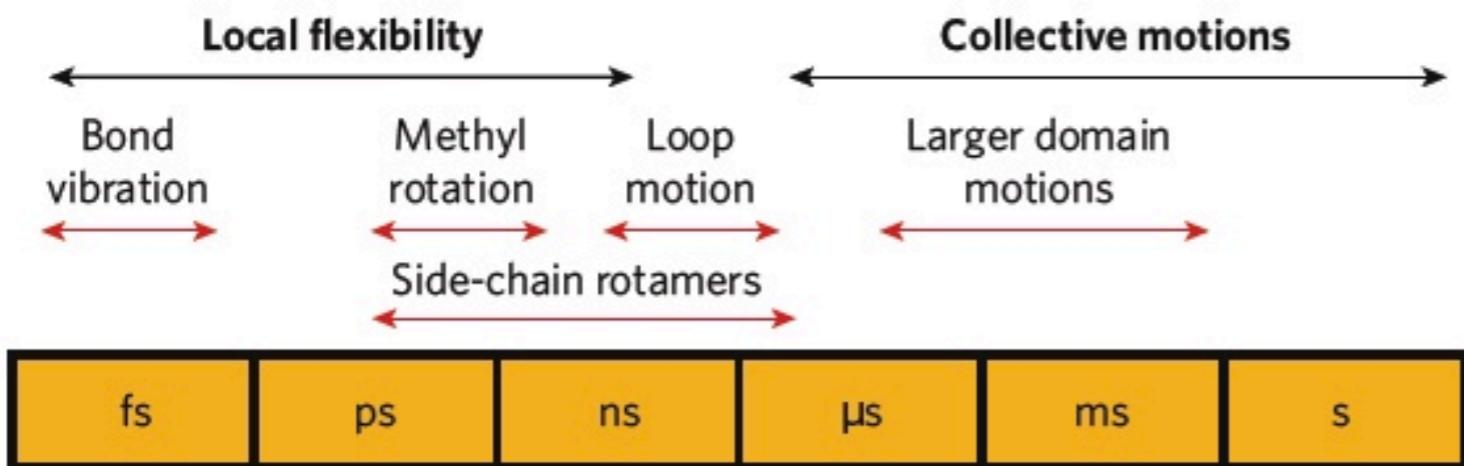
# Quantifying stoichiometries



- MS versus size-exclusion chromatography with multi-angle light scattering
- For proteins of similar composition, abundances match solution values

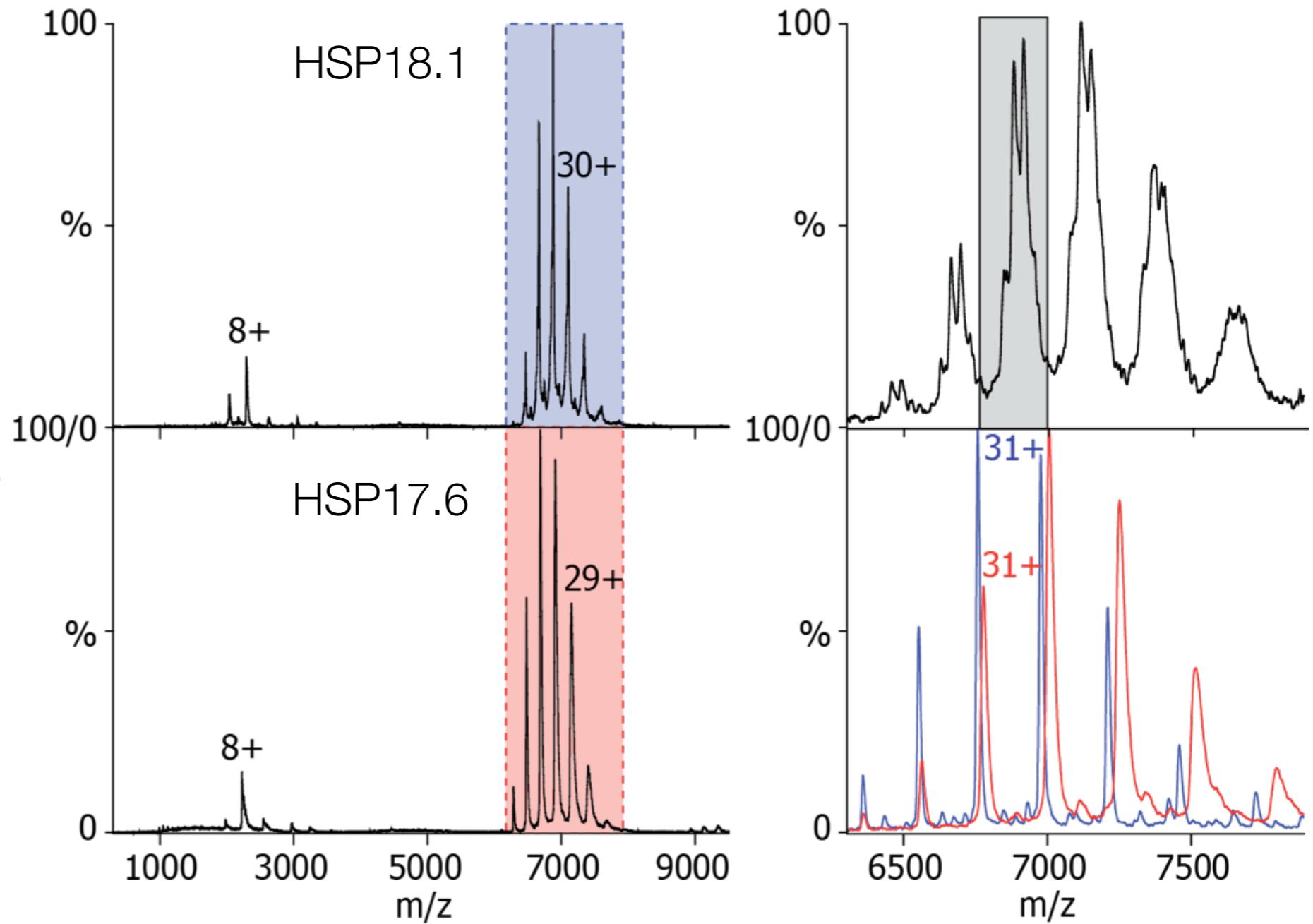
# Protein dynamics

- Proteins are not static structures, but rather undergo fluctuations both at and before equilibrium
- Such ‘protein dynamics’ are crucial to their function in the cell
- These dynamics can span a wide range of amplitudes and timescales



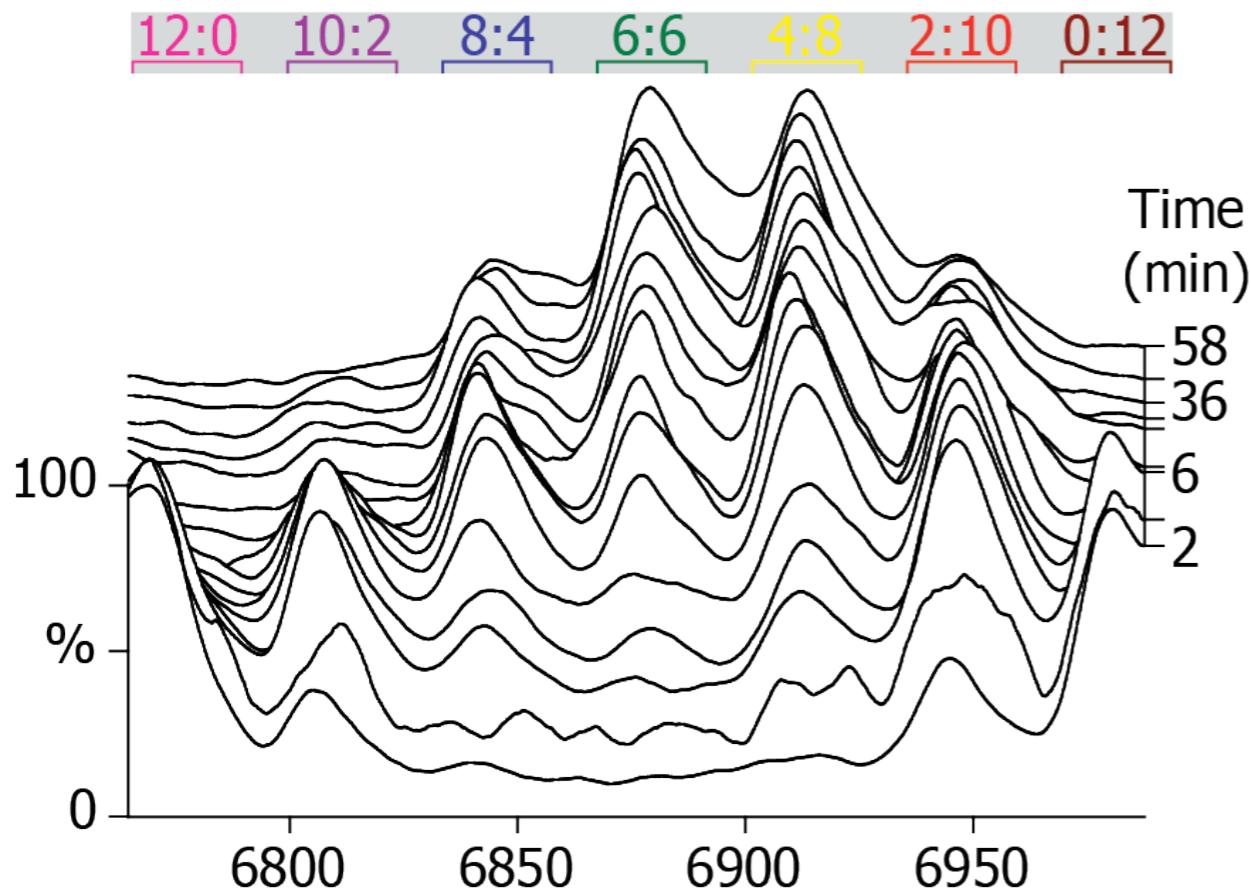
# Quaternary dynamics - Example

- Two homologous proteins from the same cellular compartment incubated
- Subunit exchange results in the appearance of hetero-oligomers

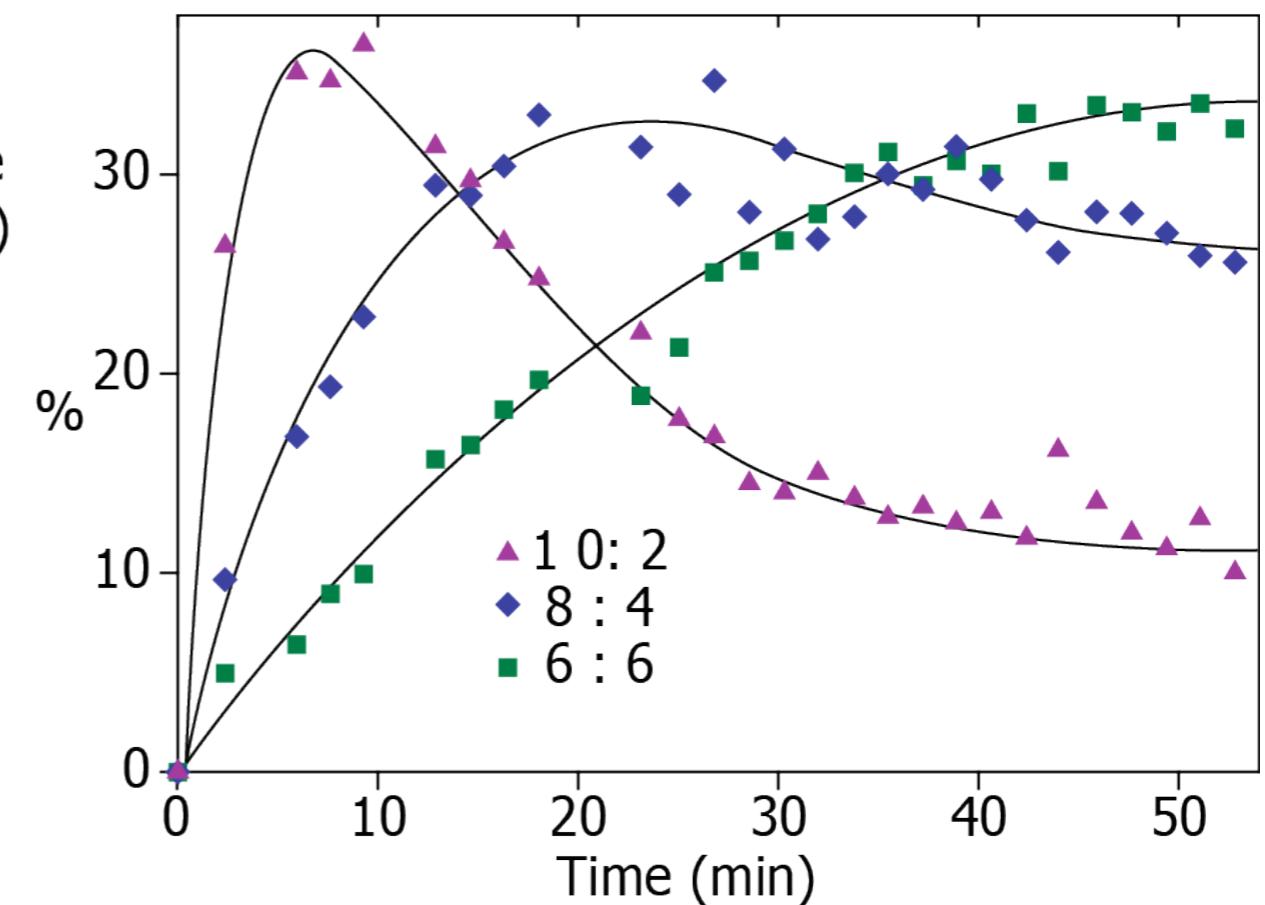


# Quaternary dynamics - Example

HSP17.6



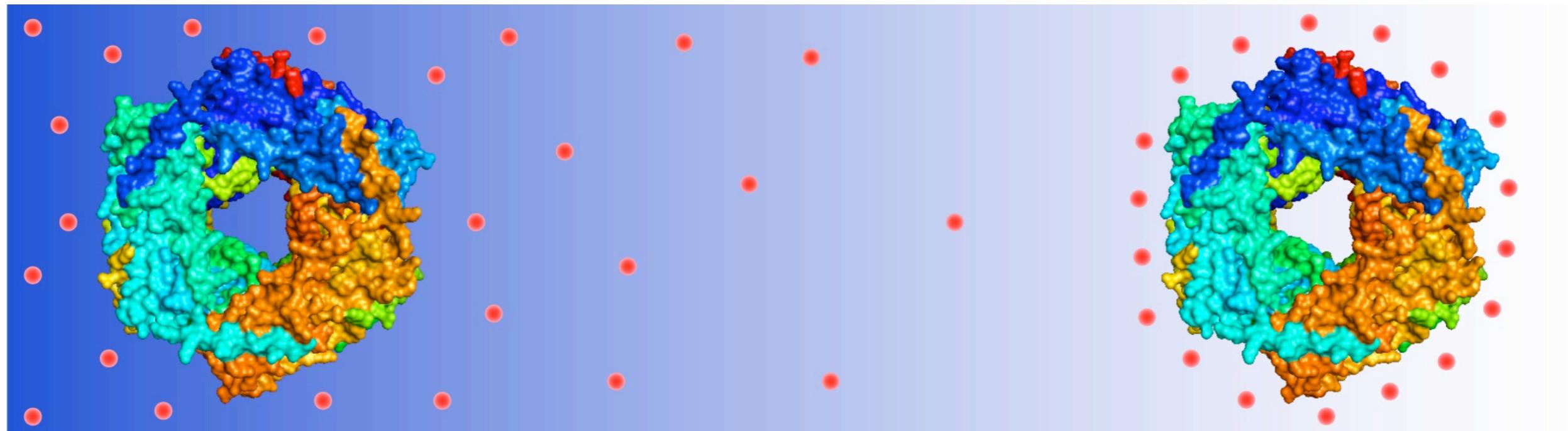
HSP18.1



- Exchange proceeds via the movement of dimeric units
- Incorporation is via sequential incorporation of dimers into oligomers
- Hetero-assembly leads to a wide variety of possible oligomers

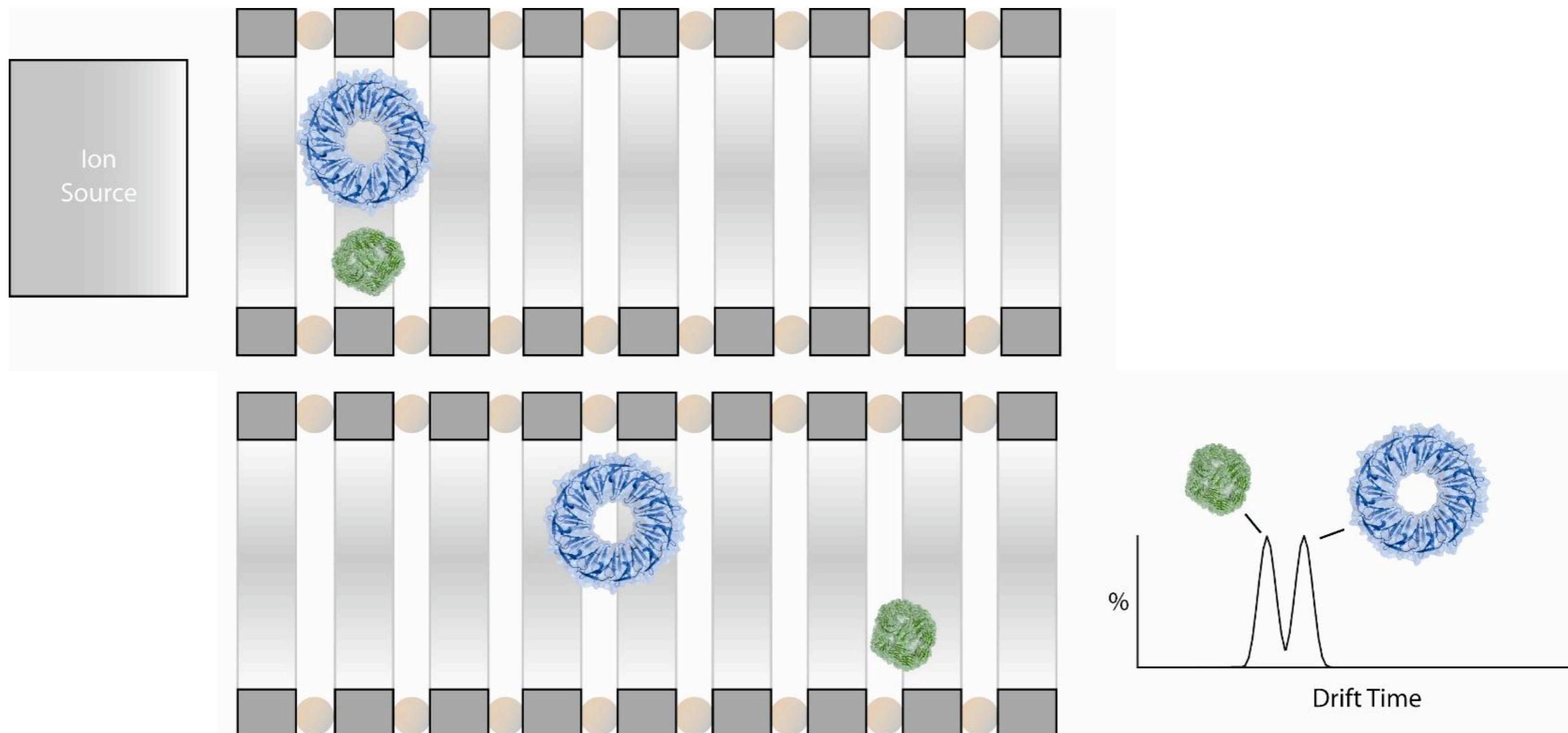
# Preservation of structure

---



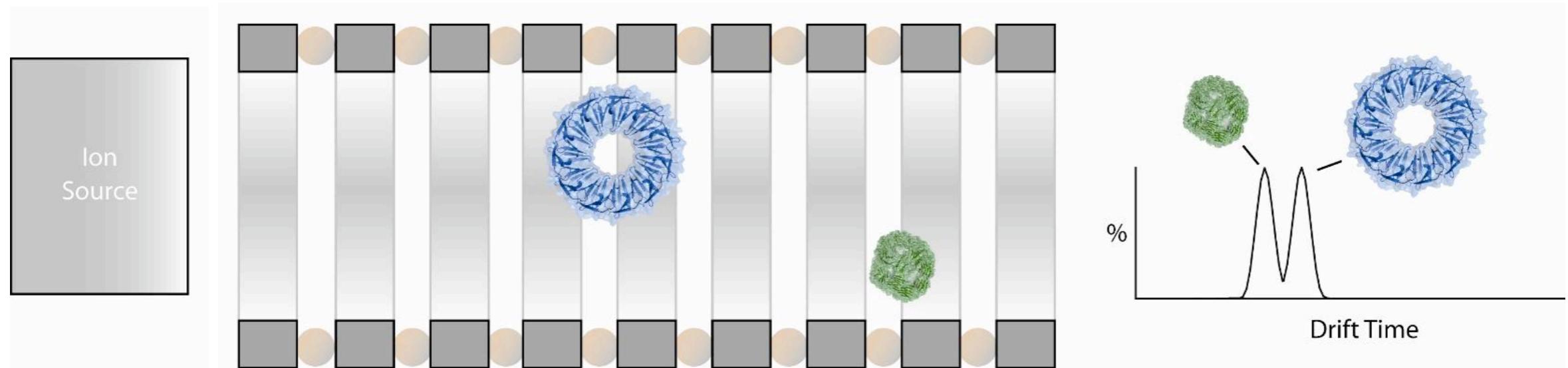
- It is clear stoichiometry is preserved in the mass spectrometer, but can we probe native structure?

# Ion mobility spectrometry (IMS)



- Separation of ions according to their ability to traverse a region of gas under the influence of a weak electric field
- Separation is based on ion ‘mobility’, unlike time-of-flight separation (mass)

# Factors contributing to IM separation



$$\Omega = \frac{(18\pi)^{1/2}}{16} \frac{ze}{(k_b T)^{1/2}} \left[ \frac{1}{m_I} + \frac{1}{m_N} \right]^{1/2} \frac{t_D E}{L} \frac{760}{P} \frac{T}{273.2} \frac{1}{N}$$

- Drift time is inversely proportional to charge
- Drift time is proportional to collision cross section (CCS,  $\Omega$ )

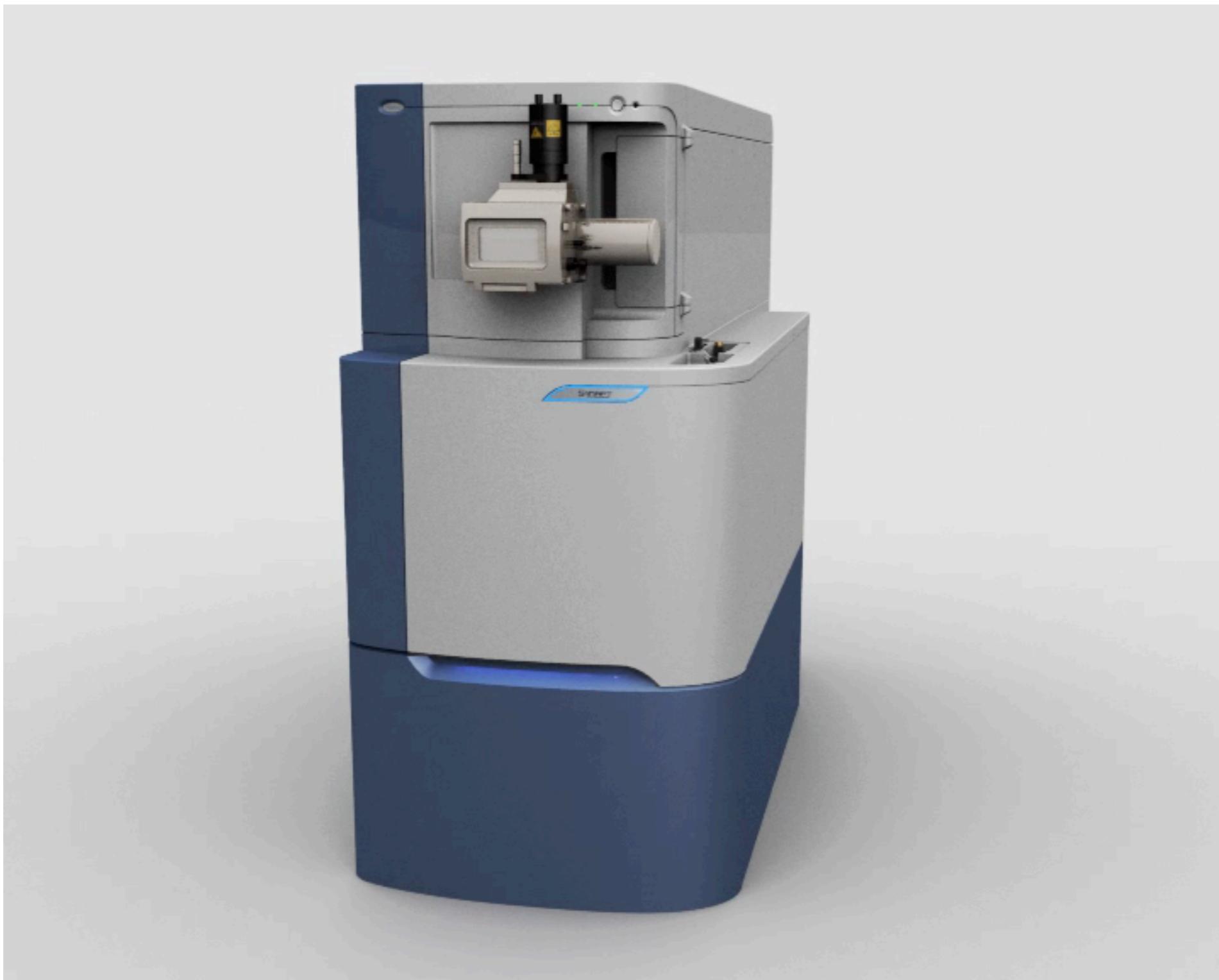
# Collision cross section (CCS)

---

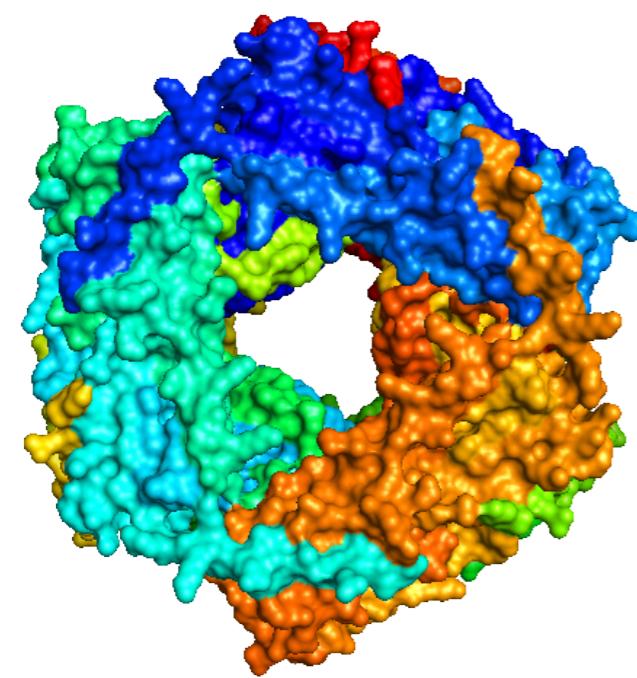
- CCS includes contribution from the gas atom
- CCS results from orientational average of analyte molecules

# IM-MS Implementation

---

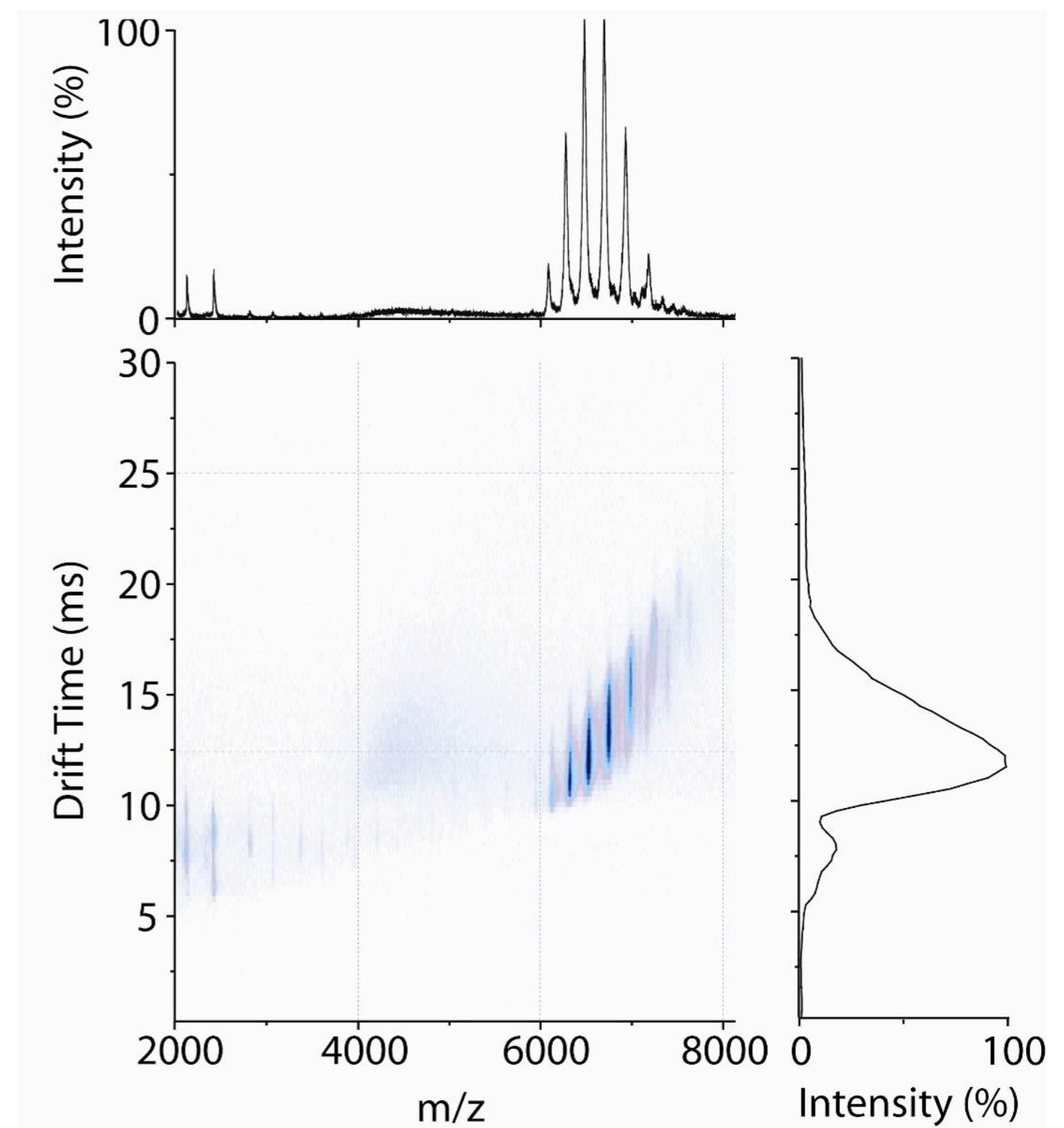


# IM-MS spectrum

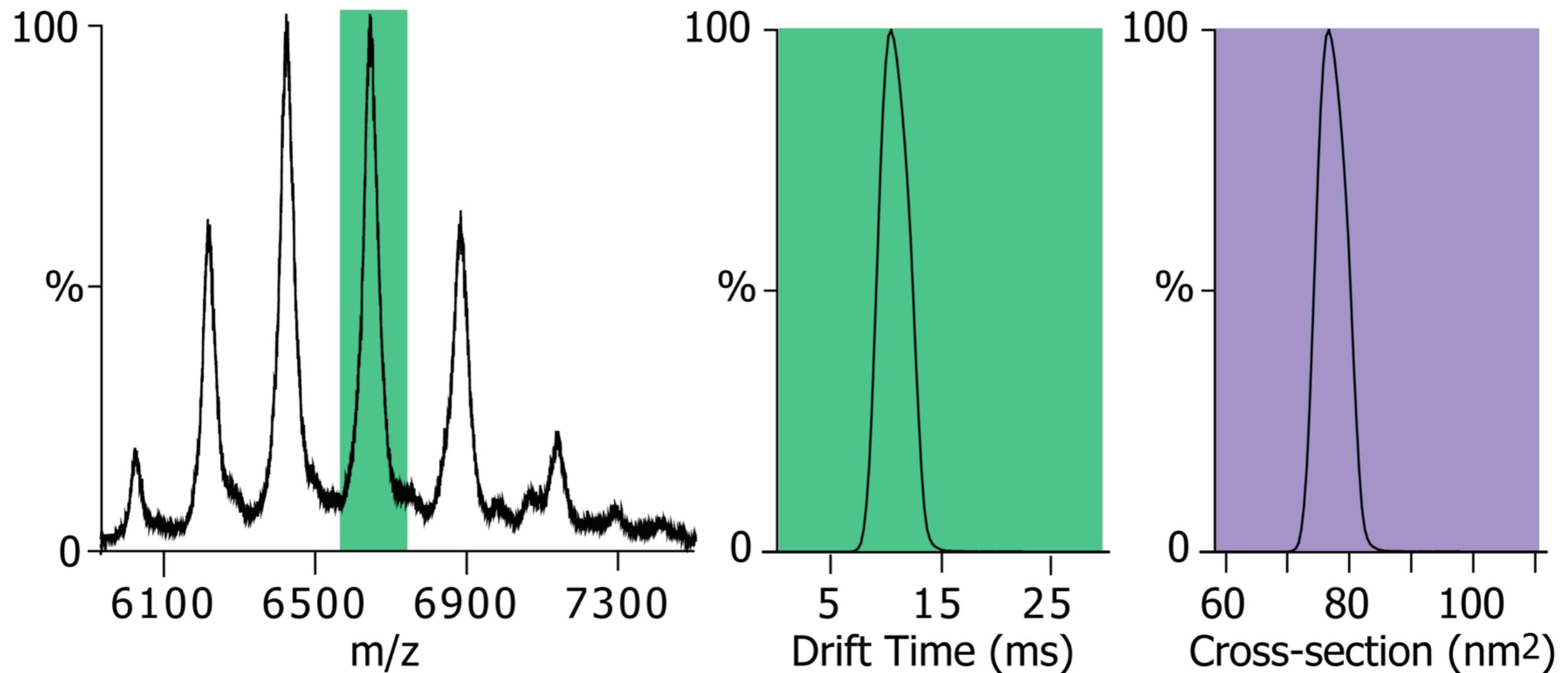


HSP16.9

- Plot of  $m/z$  versus drift time

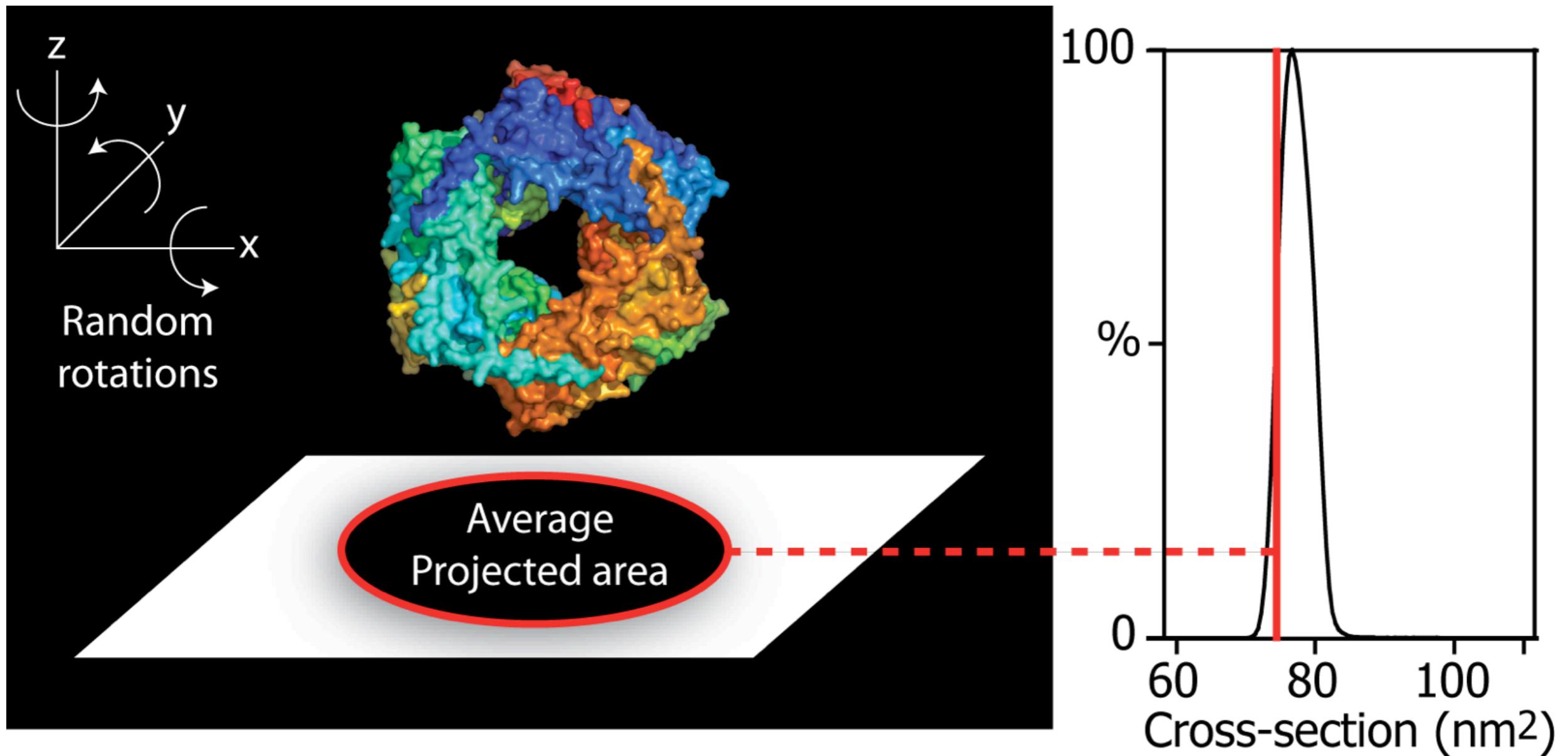


# Obtaining an experimental CCS



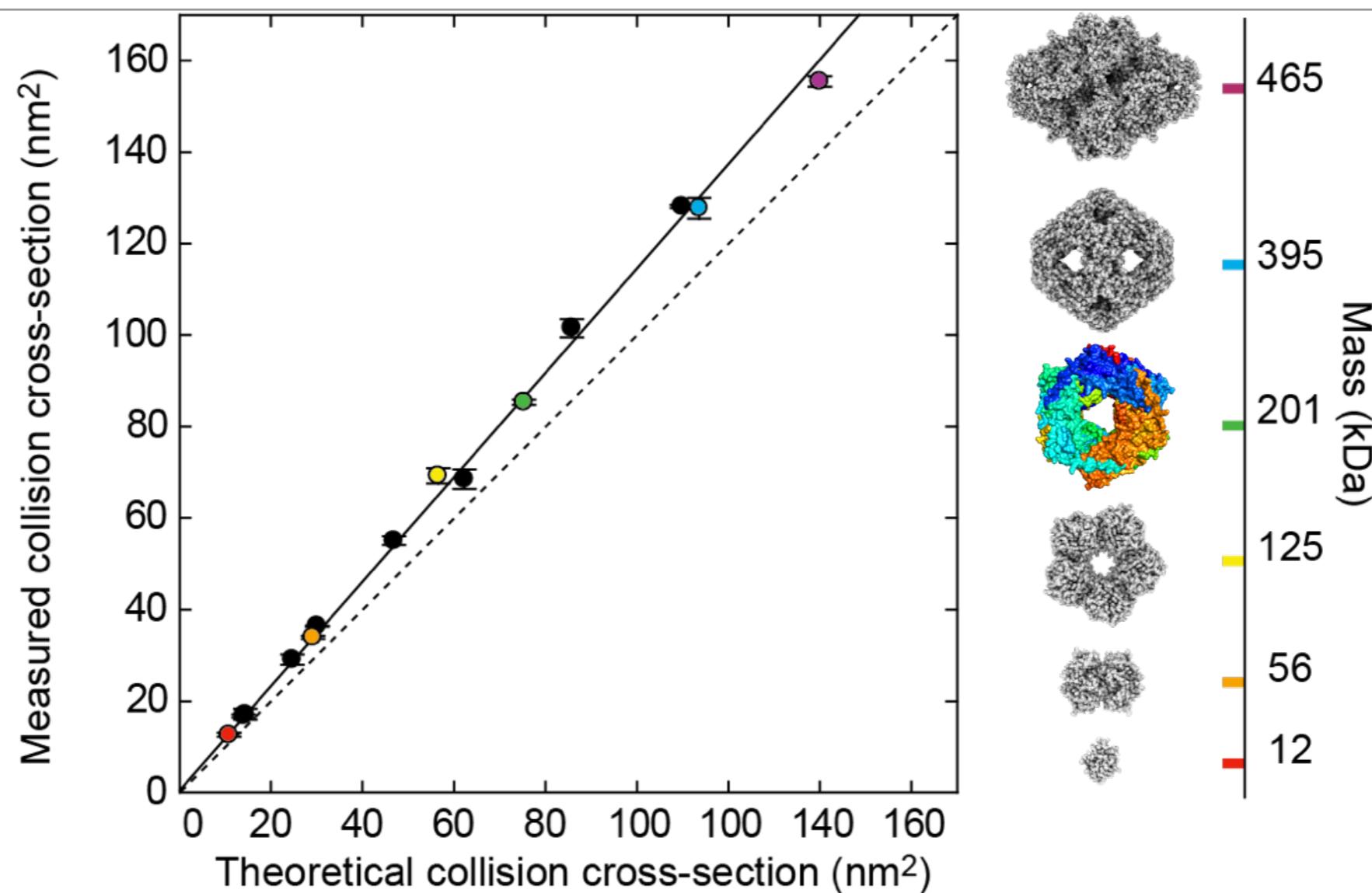
- Every feature resolved in m/z has an associated drift time distribution
- Drift time is converted into CCS either directly or via calibration

# CCS values from protein structures



- Can approximate CCS as rotationally averaged projected area
- Determine ‘theoretical’ CCSs from solved protein structures

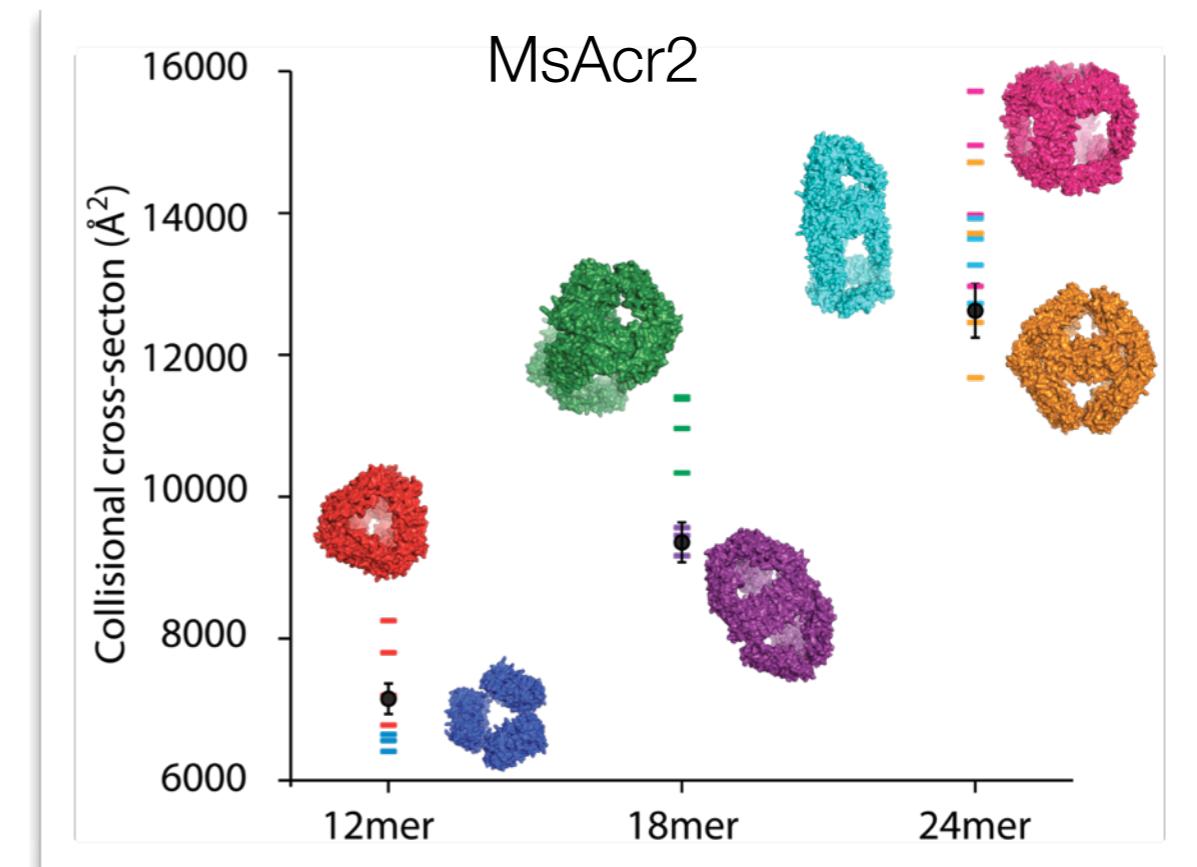
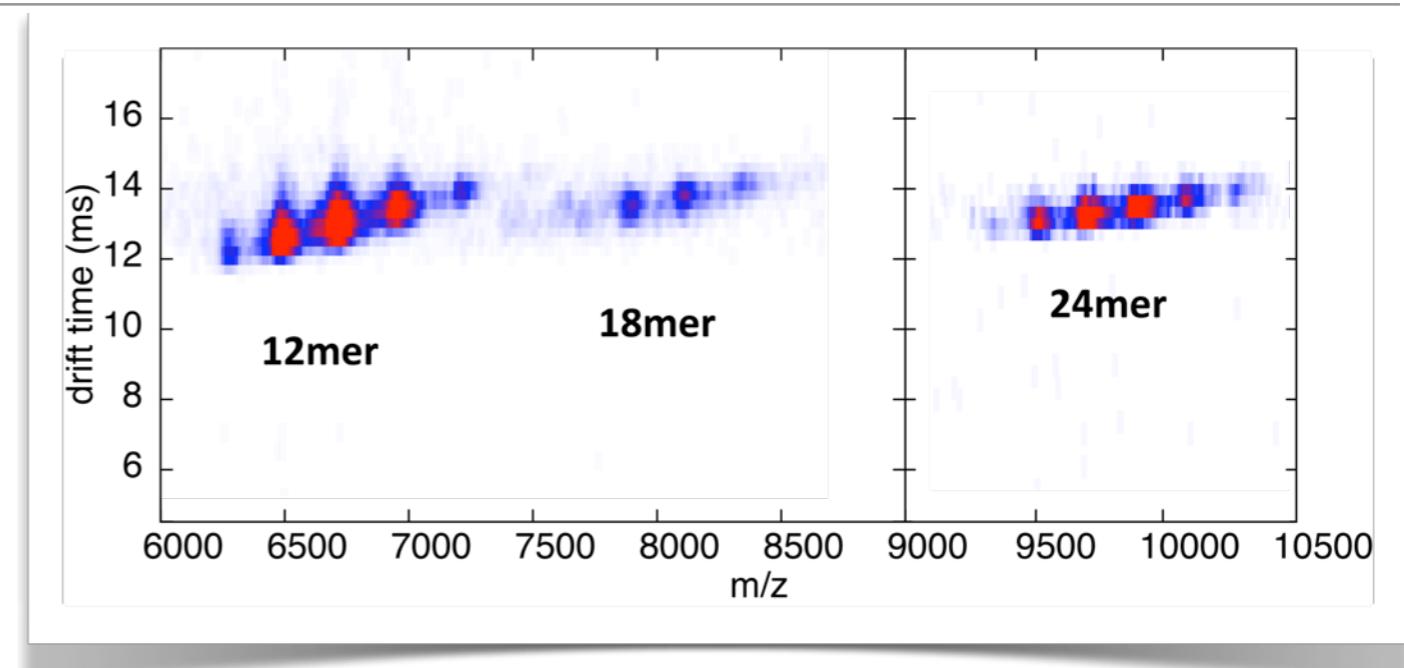
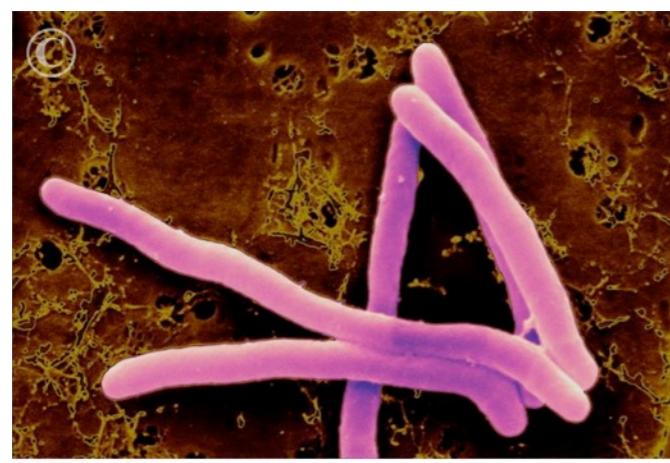
# CCS comparison



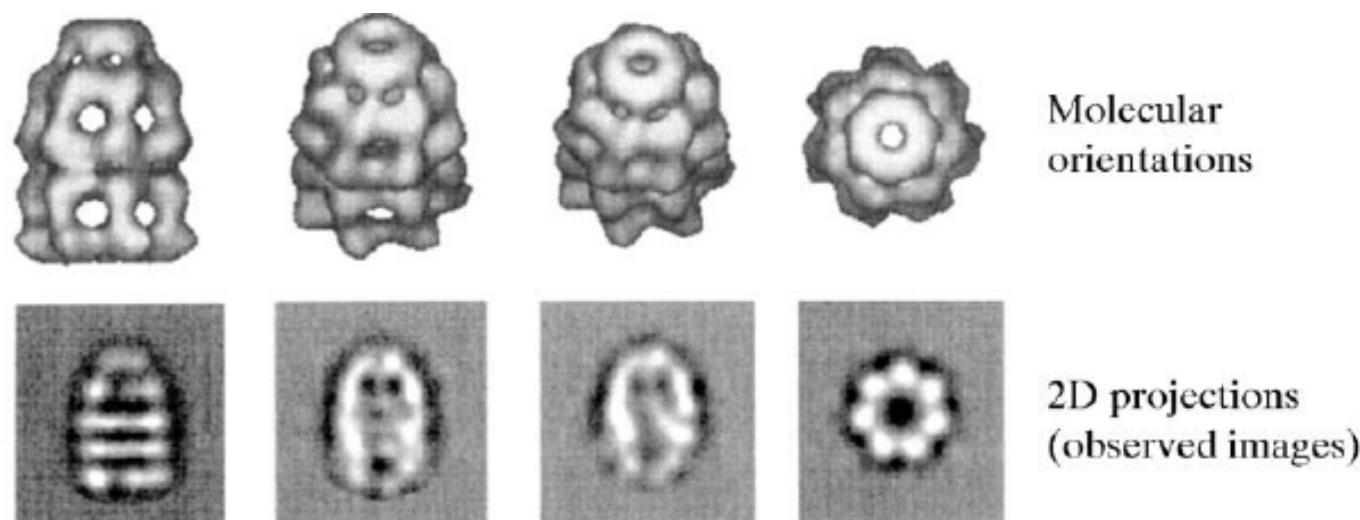
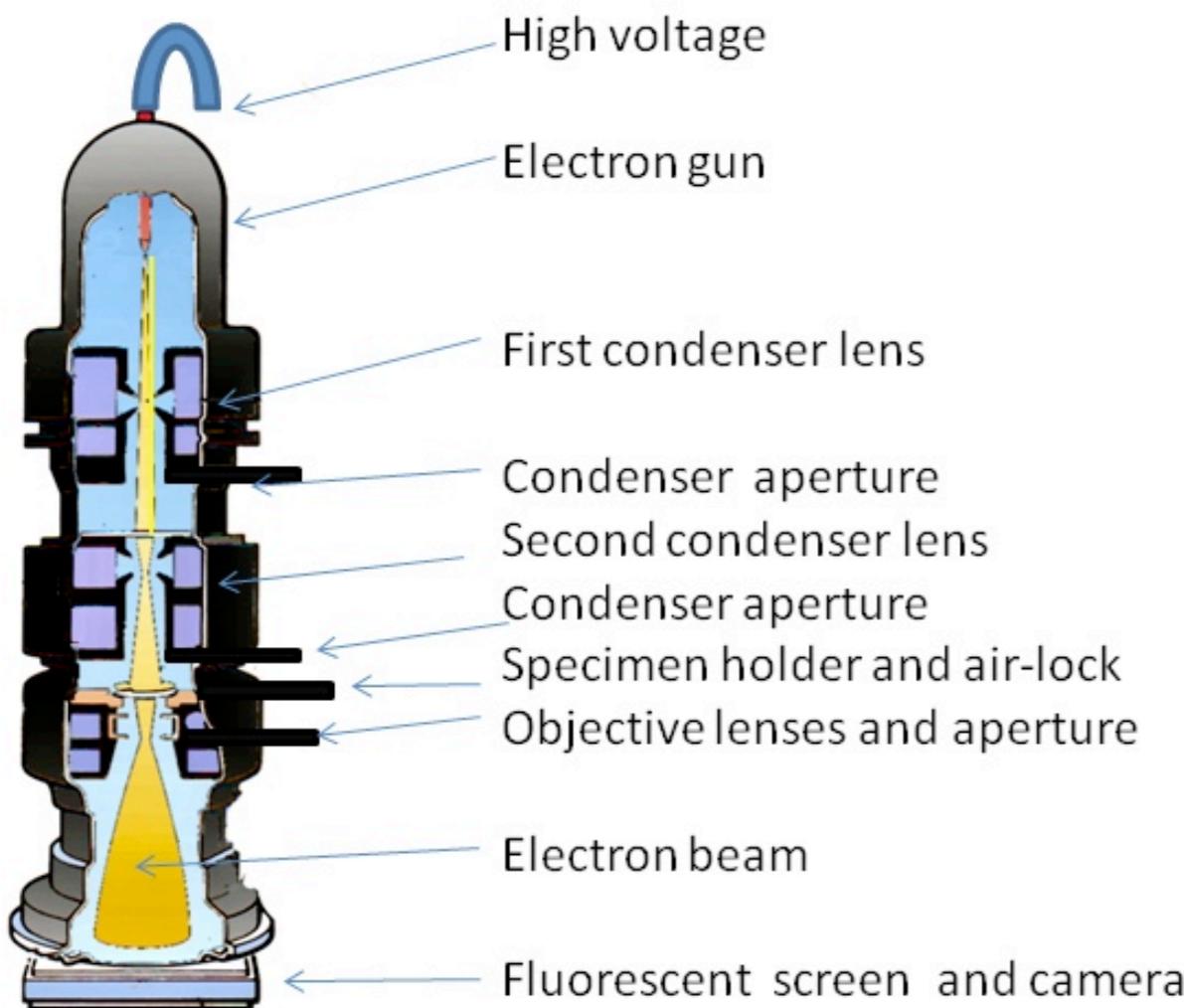
- Excellent correlation between theoretical and measured values
- Discrepancy is due to simplicity of ‘projection approximation’
- Correlation motivates use of IM measurements in assessing model structures

# Using IM-MS to filter structures - Example

- Mycobacterial small heat shock protein
- Polydisperse oligomeric protein exists in three different stoichiometries
- Based on comparison with homologous proteins likely structures are polyhedral
- Different polyhedral models can be compared to the IMS measurements

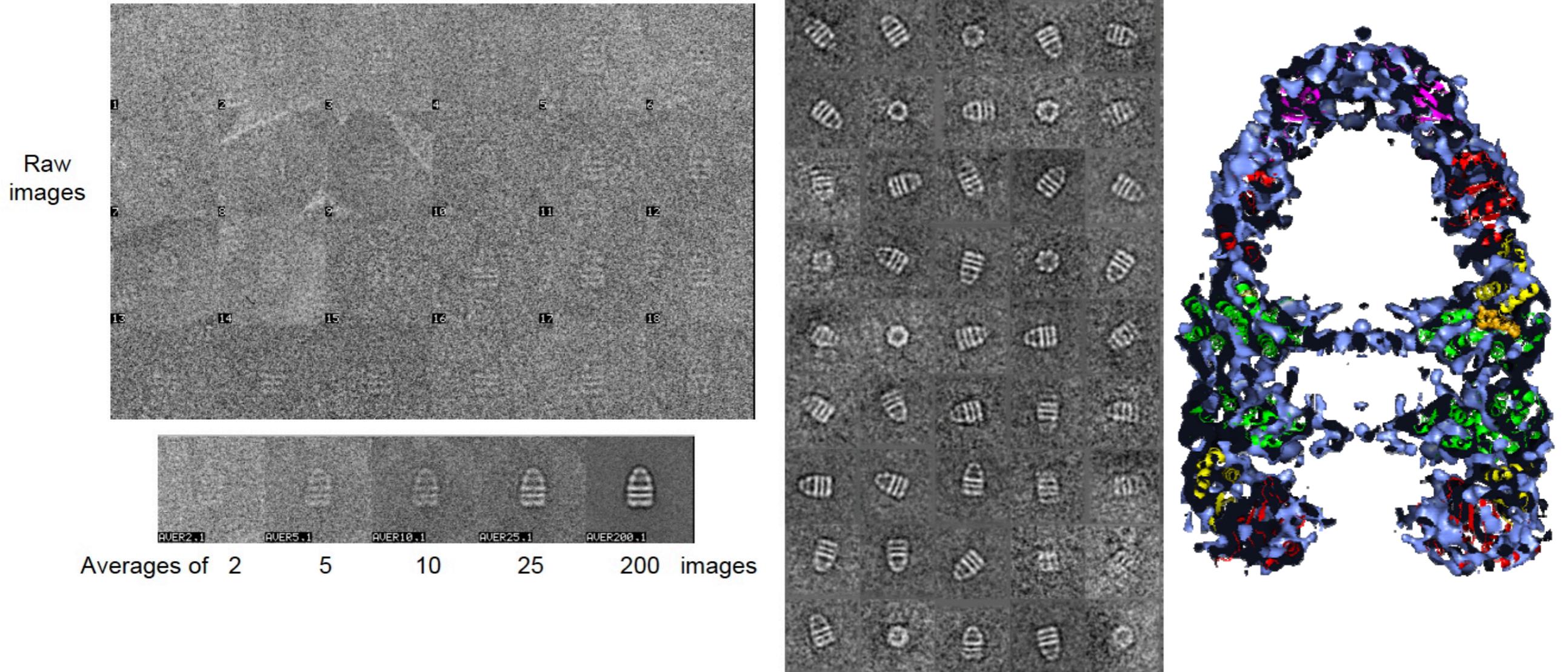


# Transmission electron microscopy (TEM)



- TEM gives two-dimensional projections of the molecular electron density

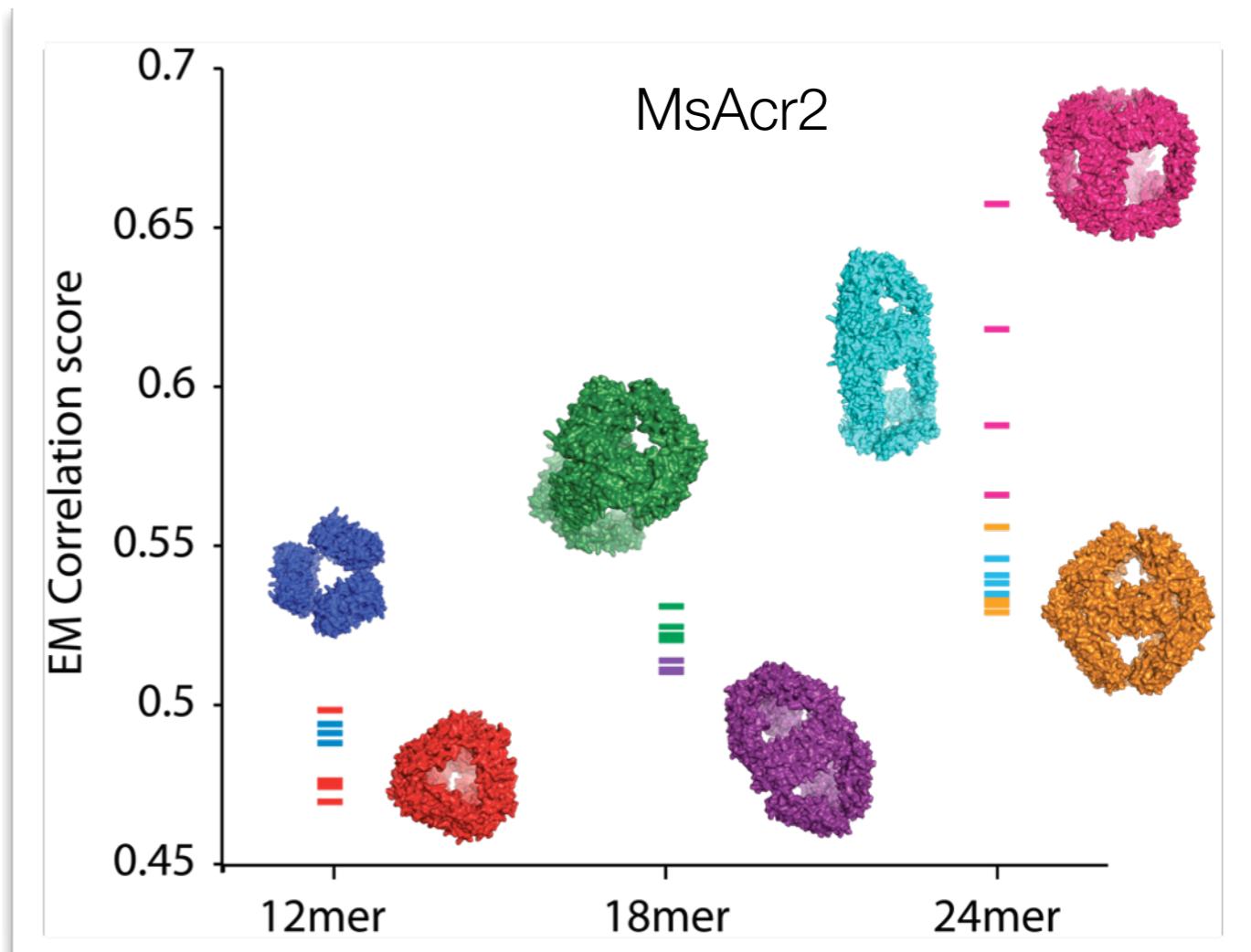
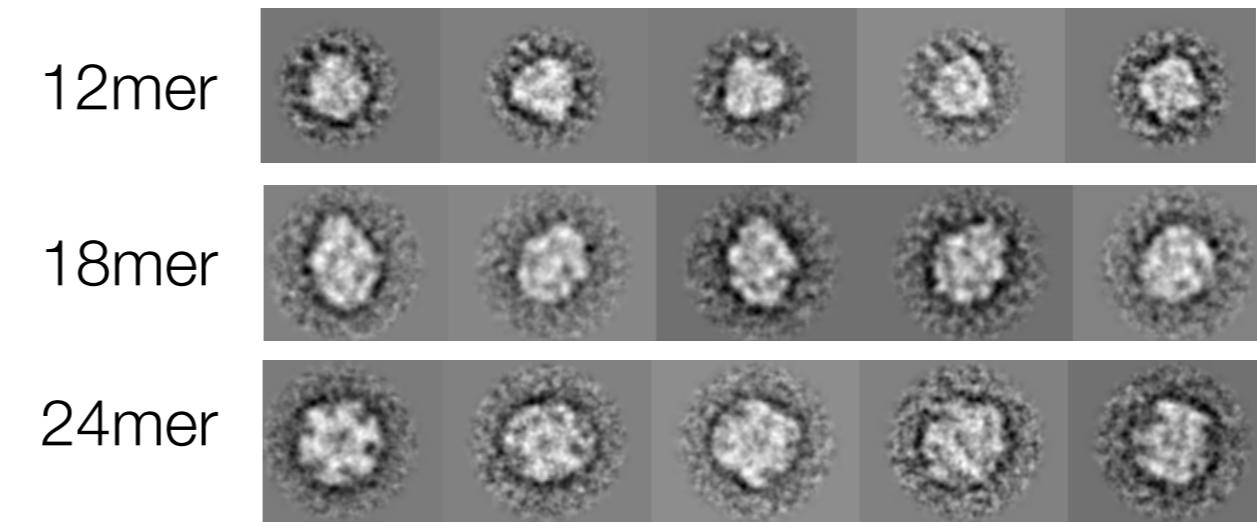
# Single particle analysis



- Align particles and average them to improve signal to noise
- Classify particles into ‘class averages’ representing particular orientations
- Combine class averages to generate 3D reconstruction

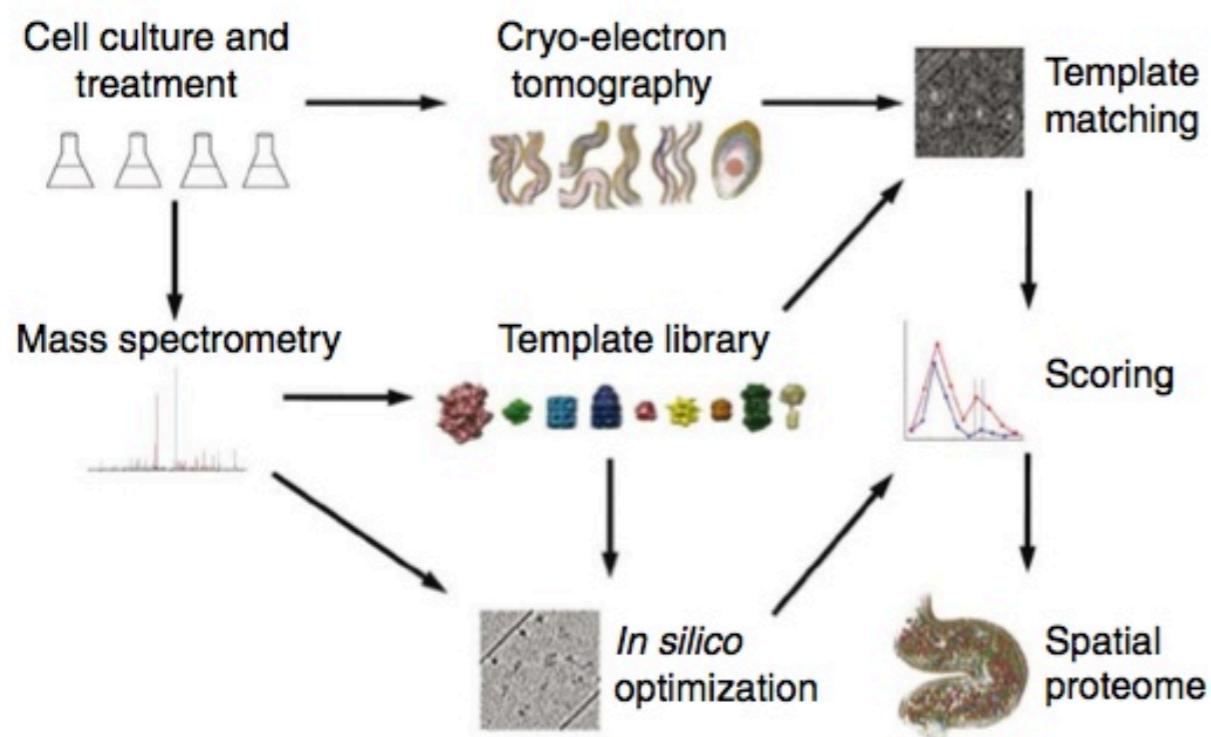
# Combining MS with TEM - Example 1

- Compare random rotations of models to TEM class averages
- Lower score is better fit
- Best fit structures are well correlated with those selected using IM-MS (see slide 27)
- Projected area from TEM is analogous to CCS area from IM

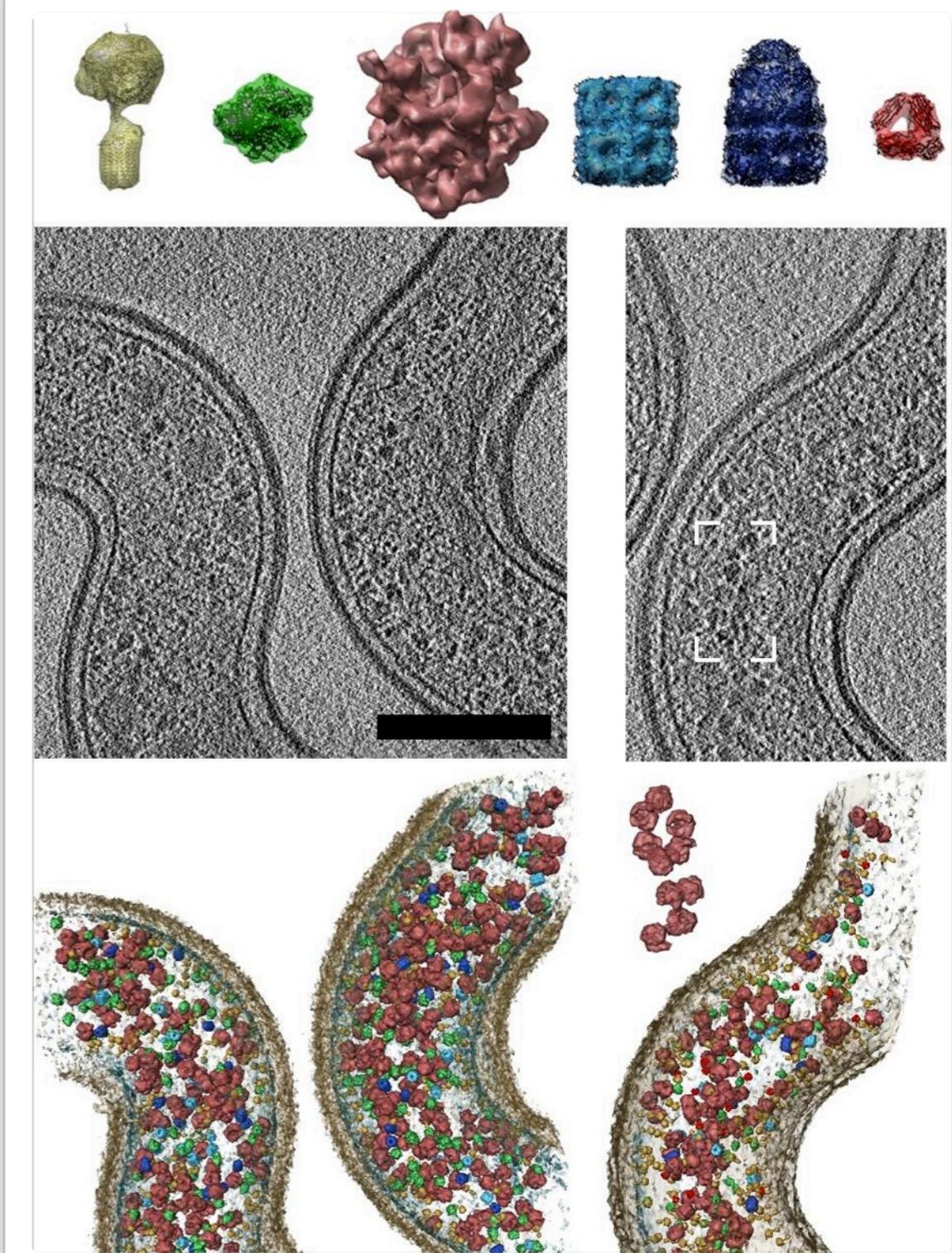


# Combining MS with EM - Example 2

- Determine abundances of proteins from MS
- Use these proteins' known structures to match to electron density in 3D reconstruction of cells
- Localise individual proteins and assess their spatial distribution



Beck ... Aebersold, Nature Methods (2009), 6, 817-23



Revision and Problem Class - Th.10, Wk1, PTCL

# Biophysical Chemistry

---

Justin Benesch  
Jonathan Doye  
Mark Wallace