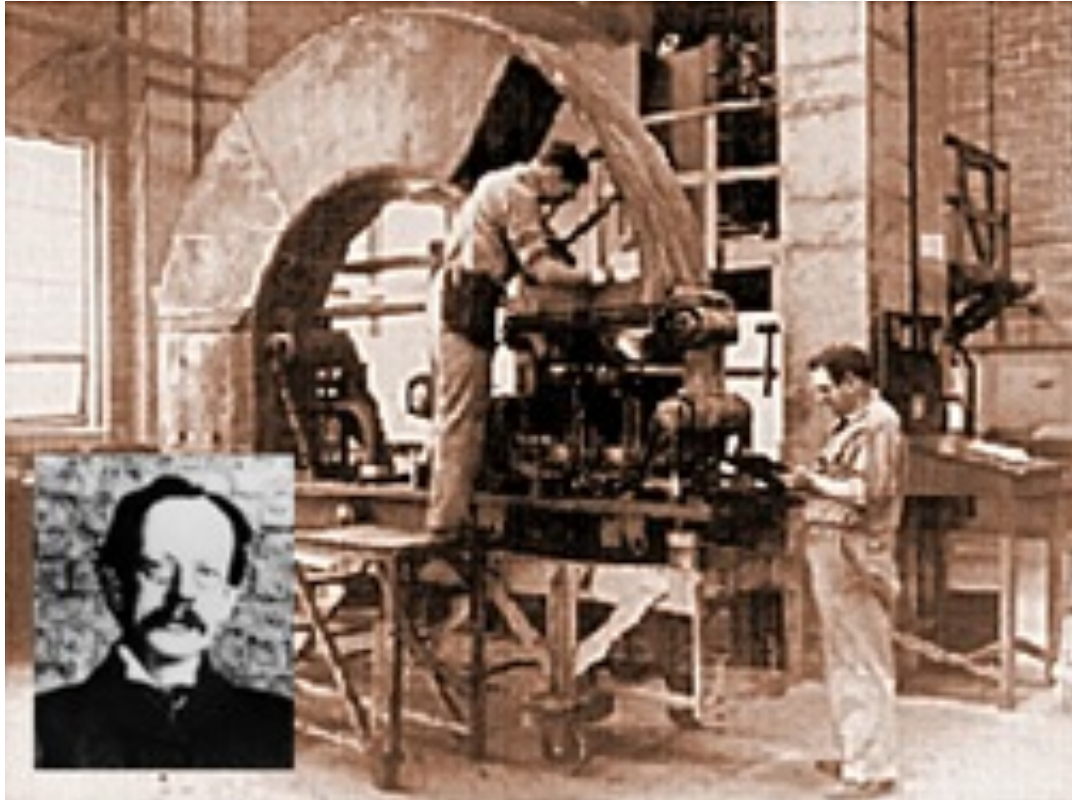
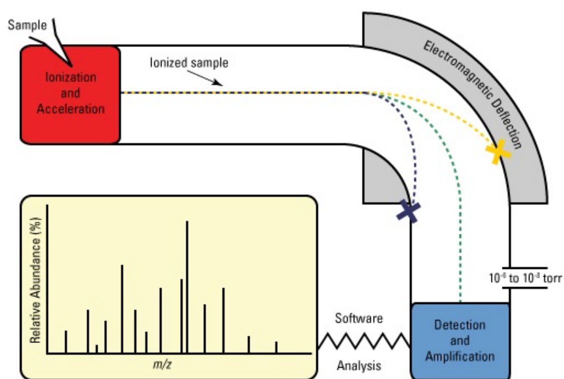
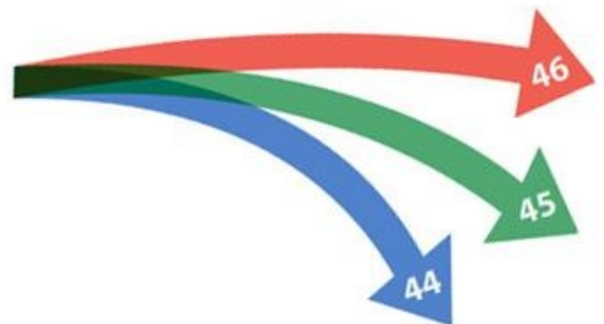
Mass spectrometry



First Concept of Time-of-Flight Mass Spectrometer, J.J Thomas April 1946





Author: Jake Ireland

Published: 29 April 2015

Chapter 1

---

1. What is mass spectrometry (MS)? what information dose MS provide.

Mass spectrometry is an analytical tool used for measuring the **molecular mass** of a sample and it can derive molecular formula. These two pieces of information are the most fundamental information you can get for a molecule. One thing mass spec can easily do is tell you about what elements are present in a molecule. This is valuable because nuclear magnetic resonance or inferred spectroscopy isn't going to give you this information in such a fundamental way.
For large samples such as biomolecules, molecular masses can be measured to high accuracy. This is sufficient to allow minor mass changes to be detected, e.g. the substitution of one amino acid for another, or a post-translational modification. To a degree it is used to confirm the nucleotide and amino acid sequences for a protein but also for more traditional small molecules.
For small organic molecules the molecular mass can be measured to within an accuracy of 5 ppm or less, which is often sufficient to confirm the molecular formula of a compound, and is also a standard requirement for publication in a chemical journal.
Structural information can be generated using certain types of mass spectrometers, usually those with multiple analysers which are known as tandem mass spectrometers. This is achieved by fragmenting the sample inside the instrument and analysing the products generated. This procedure is useful for the structural elucidation of organic compounds and for peptide or oligonucleotide sequencing.

---

2. Where is MS used?

Mass spectrometers are used in industry and academia for both routine and research purposes. The following list is just a brief summary of the major mass spectrometric applications, The ones in bold are the ones studied in this module:

- **Biotechnology: the analysis of proteins, peptides, oligonucleotides**
- **Biochemistry: Analysis of biochemicals, small molecules and large molecules.**
- Pharmaceutical: drug discovery, combinatorial chemistry, pharmacokinetics, drug metabolism
- Clinical: neonatal screening, haemoglobin analysis, drug testing
- Environmental: PAHs, PCBs, water quality, food contamination
- Geological: oil composition

---

3. How can mass spectrometry help biochemists?

- Accurate molecular weight measurements: sample confirmation, to determine the purity of a sample, to verify amino acid substitutions, to detect post-translational modifications, to calculate the number of disulphide bridges

- Reaction monitoring: to monitor enzyme reactions, chemical modification, protein digestion

- Amino acid sequencing: sequence confirmation, de novo characterisation of peptides, identification of proteins by database searching with a sequence "tag" from a proteolytic fragment

- Oligonucleotide sequencing: the characterisation or quality control of oligonucleotides (such as a primer for PCR reactions)

- Protein structure: protein folding monitored by H/D exchange, protein-ligand complex formation under physiological conditions, macromolecular structure determination

---

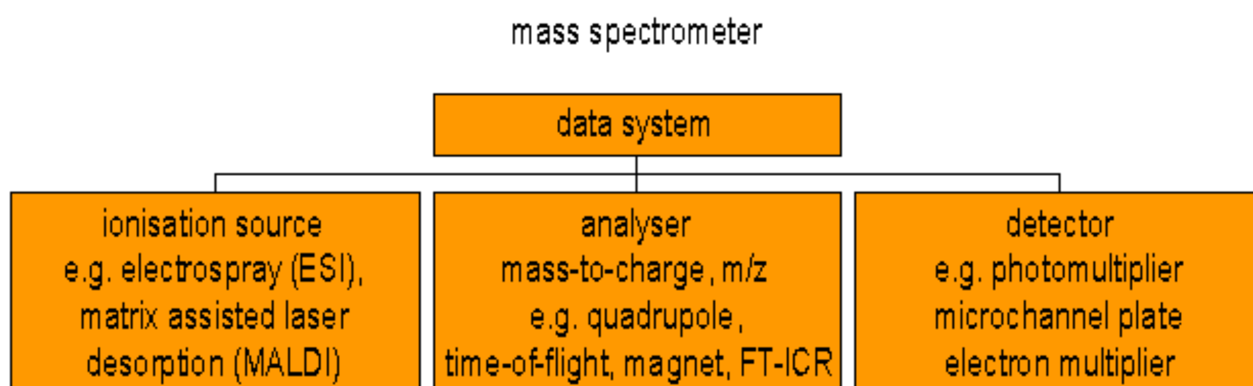
4. How dose MS work?

4.1 Introduction

Mass spectrometers can be divided into three fundamental parts, namely the ionisation source , the analyser , and the detector.

The sample has to be introduced into the ionisation source of the instrument. Once inside the ionisation source, the sample molecules are ionised, because ions are easier to manipulate than neutral molecules. These ions are extracted into the analyser region of the mass spectrometer where they are separated according to their mass (m) -to-charge (z) ratios (m/z) . The separated ions are detected and this signal sent to a data system where the m/z ratios are stored together with their relative abundance for presentation in the format of a m/z spectrum .

The analyser and detector of the mass spectrometer, and often the ionisation source too, are maintained under high vacuum to give the ions a reasonable chance of travelling from one end of the instrument to the other without any hindrance from air molecules. The entire operation of the mass spectrometer, and often the sample introduction process also, is under complete data system control on modern mass spectrometers.



Simplified schematic of a mass spectrometer

4.2 Sample introduction

The method of sample introduction to the ionisation source often depends on the ionisation method being used, as well as the type and complexity of the sample.

The sample can be inserted directly into the ionisation source, or can undergo some type of chromatography en route to the ionisation source. This latter method of sample introduction usually involves the mass spectrometer being coupled directly to a high pressure liquid chromatography (HPLC),

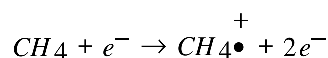
gas chromatography (GC) , and hence the sample is separated into a series of components which then enter the mass spectrometer sequentially for individual analysis.

4.3 Methods of sample ionisation

Most elements you study wont display a charge so they need to become charged. The first process developed for this was electron ionisation or (EI). this involves using electron beams to knock off electrons from a molecule.

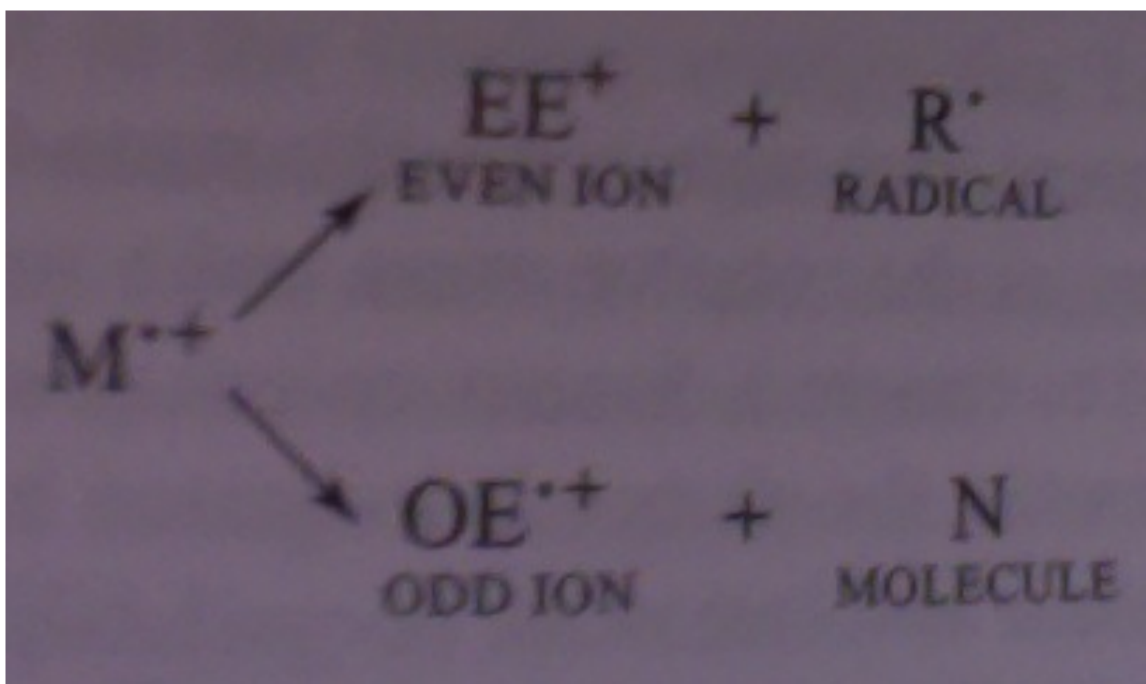
$$M + e^- + M^+ + 2e^-$$

Electrons way virtually nothing compared to molecules so they aren't counted in the MS. If you have a methyl group and knock an electron of you create a radical cation of the methyl group which is now displayed with a dot after it.



$$CH_4 + e^- \rightarrow CH_4\overset{+}{\bullet} + 2e^-$$

When you fragment an ionised molecule (a radical cation) it can happen in two ways. When you hit a molecule with an electron you introduce a lot of vibrational energy and a lot of kinetic energy and it now has a tendency to fragment. You can create an ion with an even number of electrons and a new radical species or a new neutral molecule and another radical cation.



Many ionisation methods are available and each has its own advantages and disadvantages.

The ionisation method to be used should depend on the type of sample under investigation and the mass spectrometer available.

Ionisation methods include the following:
Atmospheric Pressure Chemical Ionisation (APCI)
Chemical Ionisation (CI)
Electron Impact (EI)
Electrospray Ionisation (ESI)
Fast Atom Bombardment (FAB)
Field Desorption / Field Ionisation (FD/FI)
Matrix Assisted Laser Desorption Ionisation (MALDI)
Thermospray Ionisation (TSP)

The ionisation methods used for the majority of biochemical analyses are Electrospray Ionisation (ESI), Further analysis of these methods will come later in this paper and in weeks to come.

4.3-1 Positive and negative ionisation

With most ionisation methods there is the possibility of creating both positively and negatively charged sample ions, depending on the proton affinity of the sample. Before embarking on an analysis, the user must decide whether to detect the positively or negatively charged ions .

If the sample has functional groups that readily accept a proton (H^+) then positive ion detection is used e.g. amines $R\text{-}NH_2 + H^+ = R\text{-}NH_3^+$ as in proteins or peptides.

If the sample has functional groups that readily lose a proton then negative ion detection is used e.g. carboxylic acids $R\text{-}CO_2H = R\text{-}CO_2^-$ and alcohols R-OH = R-O- as in saccharides or oligonucleotides

4.4 Analysis and Separation of Sample Ions

The main function of the mass analyser is to separate , or resolve , the ions formed in the ionisation source of the mass spectrometer according to their mass-to-charge (m/z) ratios. There are a number of mass analysers currently available, the better known of which include quadrupoles , time-of-flight (TOF) analysers and ion traps.

These mass analysers have different features, including the m/z range that can be covered, the mass accuracy, and the achievable resolution. The compatibility of different analysers with different ionisation methods varies. For example, all of the analysers listed above can be used in conjunction with electrospray ionisation, whereas MALDI is not usually coupled to a quadrupole analyser.

Tandem (MS-MS) mass spectrometers are instruments that have more than one analyser and so can be used for structural and sequencing studies. Two, three and four analysers have all been incorporated into commercially available tandem instruments, and the analysers do not necessarily have to be of the same type, in which case the instrument is a hybrid one. More popular tandem mass spectrometers include those of the quadrupole-quadrupole, magnetic sector-quadrupole , and more recently, the quadrupole-time-of-flight geometries.

4.5 Detection and recording of sample ions.
The detector monitors the ion current, amplifies it and the signal is then transmitted to the data system where it is recorded in the form of mass spectra . The m/z values of the ions are plotted against their intensities to show the number of components in the sample, the molecular mass of each component, and the relative abundance of the various components in the sample.

The type of detector is supplied to suit the type of analyser; the more common ones are the photomultiplier , the electron multiplier and the micro-channel plate detectors.

The m/z ratio is made from arbitrary numbers but can be converted into standard units for quantitative analysis.

u = mass units $1.660540x10^{-27}kg$

e = electron charge $1.602177x10^{-19}C$

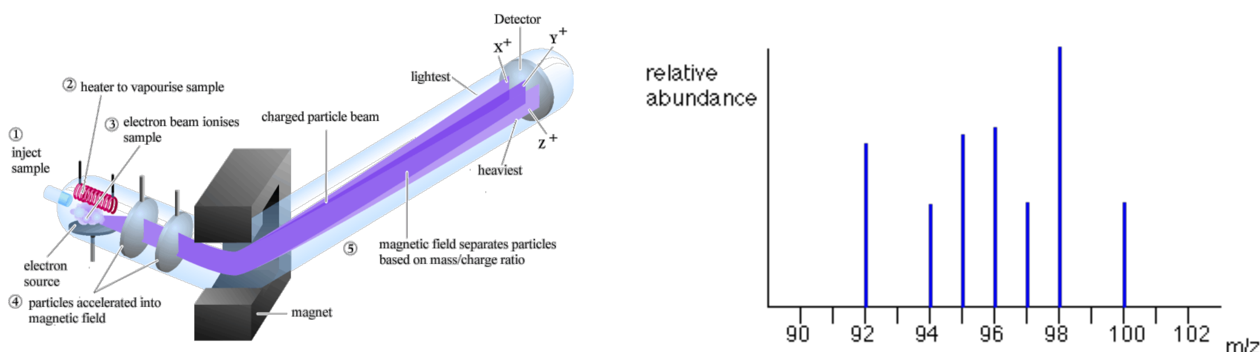
1 Thomson = 1u/e = $1.036426x10^{-8}kg\,C^{-1}$

1u = 1Da = $1.660540x10^{-27}kg$ ± 0.59ppm

In essence the mass to charge ratio is represented in mass units (u) over electron charge (e). The Thomson is the recommended ratio for the mass charge ratio where the mass 1u is equivalent to 1Da which has a degree of variation in its accuracy. This accuracy value is drawn from the most accurate average a mass detector can experimentally develop.

5.The deflection principle

A charged particle in a magnetic field is deflected. Thinking of an ionised molecule traveling on a trajectory will stay unchanged until it moves into a magnetic field.



The degree of deflection depends on the mass to charge ratio (m/z). A heavier particle is deflected less because its more massive, a smaller particle will be deflected more. A more charged particle will be deflected more than a less charged particle. Charge however comes in interregnum forms of +1, +2 etc. It cannot be 1 1/2 +. You can measure ions in the positive ion or negative ion mode but you cant do both at the same time.

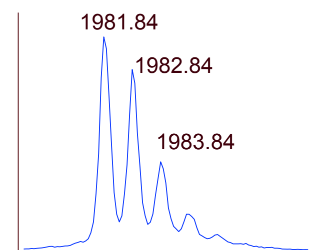
6. The nitrogen rule

when the molecule is protonated and there is an odd number of nitrogen atoms in molecule then the molecular weight will be odd. If there is an even number of nitrogens or no nitrogens present then the molecular weight will be even.

7. Isotopes

The abundance of common elements such as carbon, hydrogen, nitrogen, oxygen, sulfur, chlorine and bromine can have different isotopes that are naturally occurring in nature. These isotopes are present in different percentages. Whenmass spectrometry is used these isotopes different peaks will often reflect there natural abundances in nature.
example.

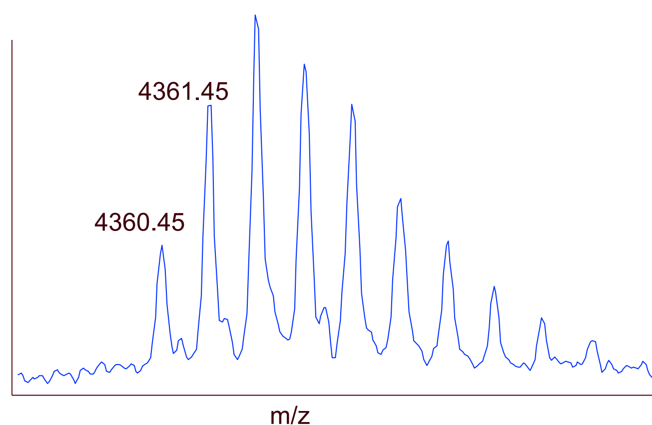
chlorine 35 is the most common isotope of clorine and will 100% natural relative abundance in a mass spectrum. Chlorine 37 is the second most abundant isotope of chlorine will have 32.5% (1/3) the relative abundance in a mass spectrum. In a compound with two chlorines the percentage of the different isotopes can be shown as to the right.



2 chlorines 35 + 35 will have 100% relative abundance peak.
2 chlorines 35 +37 will have a relative abundance (1/3*2) = 60% abundance
2 chlorines 37 +37 will have a relative abundance (1/3 * 1/3) = 11% abundance

so graphically it looks like a step down each time a new isotope is introduced. the degree to which the isotopes are naturally occurring will effect the relative abundance of the parks and thus the step down process. A compound peak that has only one type of the isotope in it is called mono-isotopic.

Larger molecules that contain a large number of the atoms will thus naturally have a similar percentage of each group of atoms. for example carbon 12 has a relative abundance of 100% while carbon 13 has a relative abundance of about 1%. For a large molecule that may contain 200 carbons you will observe different peaks for the increase possibility of carbon 13 isotopes presence. Here the 4360.45 peak is the mono-isotopic peak (two carbon 12's) and all the peaks after it have increased numbers of carbon 13 in them. statistically it is more likely to have 2 carbon 13's in a molecule with 200 carbons than it is with only carbon 12's. This is seen from the graph as the third peak is higher than the first peak.



8. Average mass

Because there a different frequencies of different isotopes of a compound in nature you cant always guarantee they you will get a mono isotopic peak. If this is coupled with poor resolution they you will get undefined peaks in your spectrogram. Resolution of a mass spectrogram is said to be the ability of the mass analyser to sperate two adjacent peaks. Seen below the difference between high resolution and poor resolution

High resolution

Poor resolution

The resolution of a mass spectrometer is the ability of it to separate adjacent peaks. when there is poor resolution it means there is poor mass accuracy. The equation to gather a resolution value is.

$$mass\ resolution = \frac{m_1}{(m_2 - m_1)}$$

When the isotopes are not resolved to a desired degree then the centre of the combined peaks for that element of compound. This average corresponds to the weighted average of all the the isotope peaks in the cluster, which is the same as the average or chemical mass. This is also referred to as the unresolved mass. The resolution ability will have a relation to the mass of the compound as the bigger the mass the less able is the mass detector able to separate peaks. Example: a 60,000Da and a 60,001Da compounds wont have separate peaks but a 6,000 and a 6,001 might have separate peaks.

9. Mass charge ratios.

The mass charge ratio, Mass to charge ratios showed the maths over there interregnums of charge. The total mass of the compound with one charge is the same as the total mass of the compound when the compound has multiple charges the masses divided by the number of charges it possesses for instance the mass of benzene is 78.046. The larger the compound is the more likely it is that you will get multiple charges and thus larger peaks for the multi charged peaks.

z	m/z	m/z
1	78.046401 / 1	78.046401
2	78.046401 / 2	39.023200
3	78.046401 / 3	26.015467

10. Types of Mass spectrometry

There are three fundamental ways of analysing a sample with a mass spectrometer.

- Full scan MS
- Tandem MS
- Accurate mass analysis

10.1 Full Scan Mass Spectrometry

This is where you place your sample into a mass spectrometer and every entity in that sample is read. The advantage of this is that you can get a wide determination of the elements and compounds present in you sample and there relative abundance to one another. The disadvantage is there will be background noises from the solution your sample may be in, there may be a large number of fragments in your sample making it hard to determine the difference and the resolution is normally not as high for an entire sample. When splitting a sample up you are able to eliminate the background noise and focus on a specific peak this is called tandem MS. this type of MS can be used to find the most abundant or least abundant peaks for selection and further analysis with tandem spectroscopy.

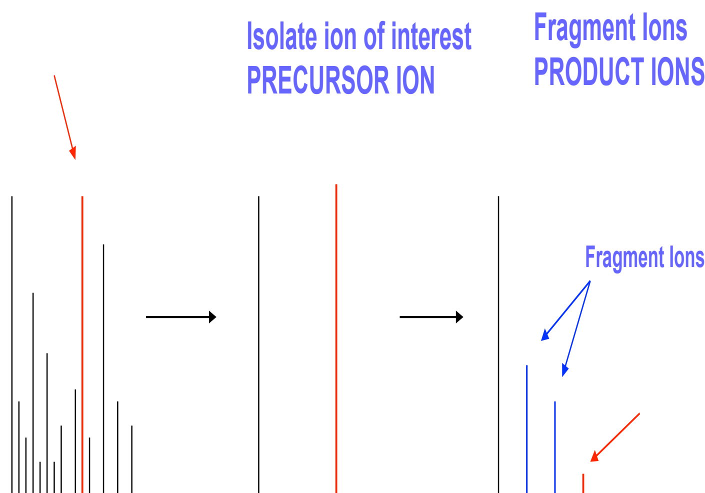
10.2 Tandem mass spectrometry (MS-MS)

MS/MS means using two mass analyzers (combined in one instrument) to select an analyte (ion) from a mixture, then generate fragments from it to give structural information of the original sample ion. Tandem mass spectrometry (MS-MS) is used to produce structural information about a compound by fragmenting specific sample ions inside the mass spectrometer and identifying the resulting fragment ions. This information can then be pieced together to generate structural information regarding the intact molecule. Tandem mass spectrometry also enables specific compounds to be detected in complex mixtures on account of their specific and characteristic fragmentation patterns.

A tandem mass spectrometer is a mass spectrometer that has more than one analyser, in practice usually two. The two analysers are separated by a collision cell into which an inert gas (e.g. argon, xenon) is admitted to collide with the selected sample ions and bring about their fragmentation. The analysers can be of the same or of different types, the most common combinations being:

- quadrupole - quadrupole.
- quadrupole - time-of-flight.

The basic modes of data acquisition for tandem mass spectrometry experiments are as follows:

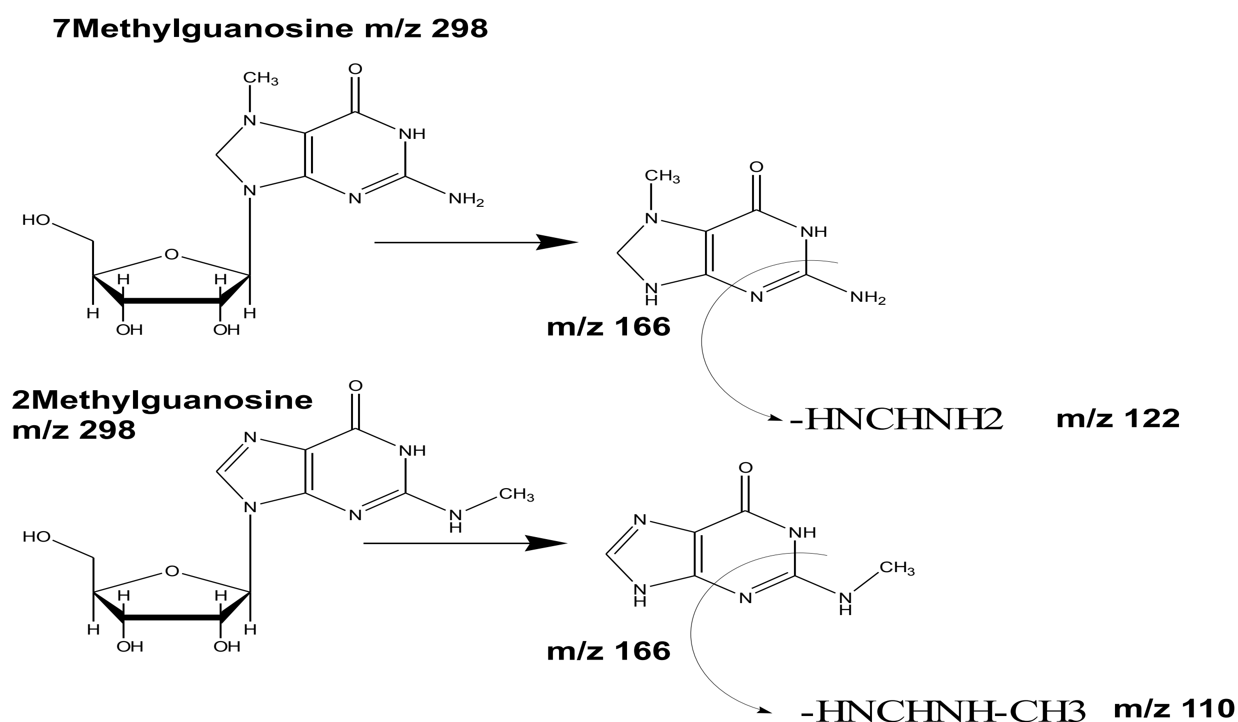


Product or daughter ion scanning:
the first analyser is used to select user-specified sample ions arising from a particular component; usually the molecular being investigated (i.e. $(M+H)^+$ or $(M-H)^-$) ions. These chosen ions pass into the collision cell, are bombarded by the gas molecules which cause fragment ions to be formed, and these fragment ions are analysed i.e. separated according to their mass to charge ratios, by the second analyser. All the fragment ions arise directly from the precursor ions specified in the experiment, and thus produce a fingerprint pattern specific to the compound under investigation.This type of experiment is particularly useful for providing structural information concerning small organic molecules and for generating

peptide sequence information.

Precursor or parent ion scanning:
the first analyser allows the transmission of all sample ions, whilst the second analyser is set to monitor specific fragment ions, which are generated by bombardment of the sample ions with the collision gas in the collision cell. This type of experiment is particularly useful for monitoring groups of compounds contained within a mixture which fragment to produce common fragment ions.

This type of analysis gives you the power to isolate specific compounds and determine there structural makeup like this example with methyl-guanosine. There m/z peaks will be the same (exactly the same) because they have exactly the same mass. However further fragmentation of this compound can determine where the methyl is from the two possible locations.



Selected/multiple reaction monitoring:
both of the analysers are static in this case as user-selected specific ions are transmitted through the first analyser and user-selected specific fragments arising from these ions are measured by the second analyser. The compound under scrutiny must be known and have been well-characterised previously before this type of experiment is undertaken. This methodology is used to confirm unambiguously the presence of a compound in a matrix e.g. drug testing with blood or urine samples. It is not only a highly specific method but also has very high sensitivity.

10.3 accurate mass analysis.

Some elemental formulas can have the same mass but totally different atoms. An example of this is oxygen16 and methane C12 + 4H4's. Accurate mass analysis is a type of mass detection that increases the sensitivity. The mass will normally be given in ppm and by increasing the sensitivity you increase the accuracy of the ppm of your sample. For example the more sensitive a detector is the grater number of decimal places you can get. you generally need

<5ppm accuracy, ideally 3.5ppm, to determine an elemental formula. This correlates to 3decimal places which for the difference between methan and oxygen is 0.04306.

H = 1.00794
C = 12.0107
O = 15.9994

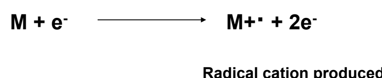
Methane = 12.0107 + (4 x 1.00794) = 16.04246
Oxygen = 15.9994

The difference between them is 0.04306

11. Types of ionisation source

Electron ionisation (EI)

This involves the analyte to pass through an electrical field. The electrons do not interact with the compound but they do disrupt bonds within the compound and encourage electrons to dissociate from the compound. This produces positive ions and has a very low rate of success. There is a 1 - 1000 compound ionisation which is poor for very low concentrations of a sample. The sample must be in gas phase before it passes through the election beam



Chemical Ionisation

chemical ionisation is sometimes referred to as soft ionisation as it is not as disrupting to a molecule as an electrical field. ionisation occurs through the collision of the analyte with ions of a reagent gas that are present in the ion source. Some common reagent gases include: methane, ammonia, and isobutane. Inside the ion source, the reagent gas is present in large excess compared to the analyte. The molecular weight limit of the compounds being ionised is 1000Da and the compounds must be volatile.

Fast atom bombardment

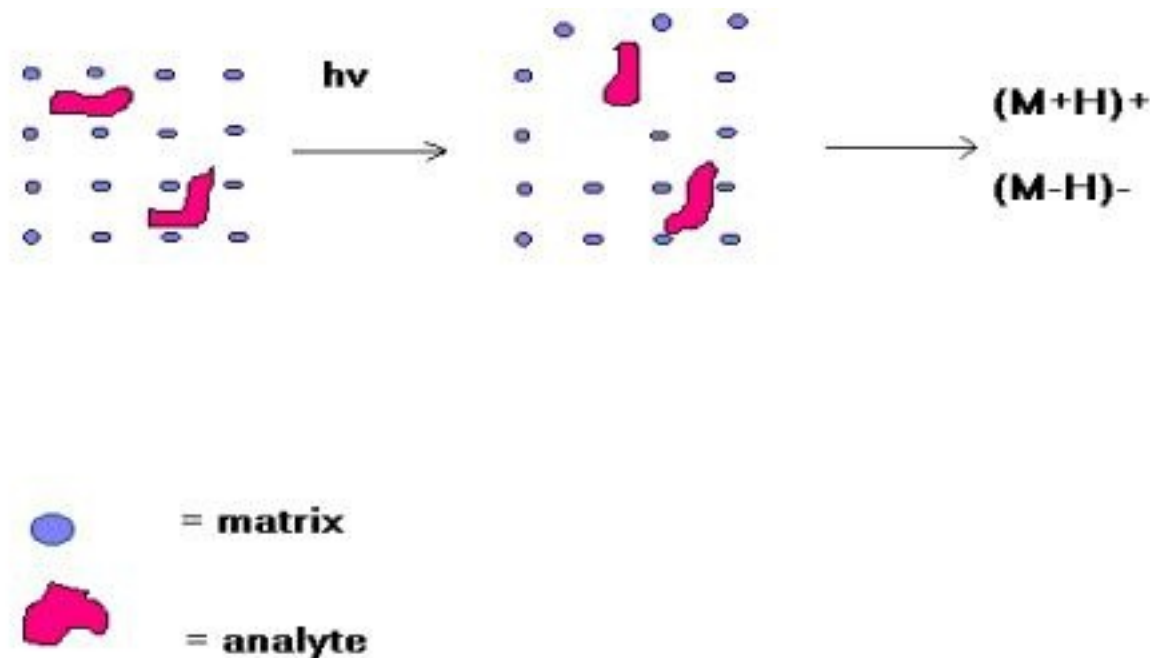
The material to be analyzed is mixed with a non-volatile chemical protection environment called a matrix (usually glycerol) and is bombarded under vacuum with a high energy (4000 to 10,000 electron volts) beam of atoms. The atoms are typically from an inert gas such as argon or xenon. FAB is a relatively soft ionization technique and produces primarily intact protonated molecules denoted as $[M+H]^+$ and deprotonated molecules such as $[M-H]^-$. There is a usually peaks in the spectrogram from every component of the matrices fragmentation but there is a also very little fragmentation seen from the analytes.

12. MALDI (matrix assisted laser dispersion ionisation)

Matrix Assisted Laser Desorption Ionisation (MALDI) deals well with thermolabile, non-volatile organic compounds especially those of high molecular mass and is used successfully in biochemical areas for the analysis of proteins, peptides, glycoproteins, oligosaccharides, and oligonucleotides. It is relatively straightforward to use and reasonably tolerant to buffers and other additives. The mass accuracy depends on the type and performance of the analyser of the mass spectrometer, but most modern instruments should be capable of measuring masses to within 0.01% of the molecular mass of the sample, at least up to 40,000 Da.

MALDI is based on the bombardment of sample molecules with a laser light to bring about sample ionisation. The sample is pre-mixed with a highly absorbing matrix compound for the most consistent and reliable results, and a low

concentration of sample to matrix works best. The matrix transforms the laser energy into excitation energy for the sample, which leads to sputtering of analyte and matrix ions from the surface of the mixture. In this way energy transfer is efficient and also the analyte molecules are spared excessive direct energy that may otherwise cause decomposition. Most commercially available MALDI mass spectrometers now have a pulsed nitrogen laser of wavelength 337 nm.



Matrix assisted laser desorption ionisation (MALDI)

The sample to be analysed is dissolved in an appropriate volatile solvent, usually with a trace of trifluoroacetic acid if positive ionisation is being used. A range of compounds is suitable for use as matrices: sinapinic acid is a common one for protein analysis while alpha-cyano-4-hydroxycinnamic acid is often used for peptide analysis. The final solution is applied to the sample target which is allowed to dry prior to insertion into the high vacuum of the mass spectrometer. The laser is fired, the energy arriving at the sample/matrix surface optimised, and data accumulated until a m/z spectrum of reasonable intensity has been amassed. The time-of-flight analyser separates ions according to their mass(m)-to-charge(z) (m/z) ratios by measuring the time it takes for ions to travel through a field free region known as the flight, or drift, tube. The heavier ions are slower than the lighter ones.

The m/z scale of the mass spectrometer is calibrated with a known sample that can either be analysed independently (external calibration) or pre-mixed with the sample and matrix (internal calibration).

The Laser focused to a 30-50µm spot size and The matrix absorbs the energy of the laser. Because the laser is a pulse laser typically 10-100 lazer shots are accumulated. This improves the shot to particle ratio. The Large amount of energy – localised laser heating causes a sublimation of the matrix analyte mix. you can sometimes creat a spot that will have an unequal sharing of the analyse across the spot. Shooting the laser at a number of random locations within the spot you can average the analyse readings.

There is a minor disadvantage to the use of the MALDI system over the gas chromatography. This comes from the analyte being dissolved in solvent. The mixing of the analyse and the solvent is called a matrix and this is dried on to matrix plate. This dried matrix essentially becomes a salt deposit on the matrix slide. This matrix layer is bombarded with short laser pulses that will cause localised matrix sublimation. The matrix solution is made from small acids that will contain a benzene ring capable of absorbing the lasers energy at the given wavelength. This protects the analyte from laser bombardment and subsequent fractionation. The matrixes rapid heating causes an expansion into a gaseous phase that will contain the analyte. The disadvantage of this is that the matrix molecules will enter the time

of flight tube with the analyte. It is possible to partially separate these two products, both before they enter the tube and after. Before with mass filters such as quadrupoles, and after with the time of flight method where smaller molecules will travel faster than larger molecules. The MALDI spectra will however still contain the peaks from the matrix acids though. These acid matrix peaks may influence the same peak reading as the fractions of the analyte and create inaccurate indications of how much analyte is present. This is considered an interference with the analyse analysis and can lead to obscure readings (Hoffmann, Stroobant, 2007, p33-41). MALDI-TOF MS was essentially developed to analyse compounds of large molecular masses so that they don't interfere with the molecular matrix peaks and can thus be clearly distinguished. The lower the molecular masses your compound gets the grater the possibility of the matrix interfering with the product peaks. The lowest recommended molecular weight of a analyte is 500Da. The common oleic acid fatty acid C18 has a mass of 282.4Da which obviously falls below the recommended limit. A solution to this problem could come from putting the fatty acids in a solution of of much larger molecular weight like meso-tetrakis(pentaflurophynl)-porphyrin (figure 1) which docent create peaks at the lower m/z values like fatty acids (Hillenkamp, Peter-Katalinić, 2007, p218-238) (Ayorinde, Hambright, Porter and Keith, Jr, 1999)

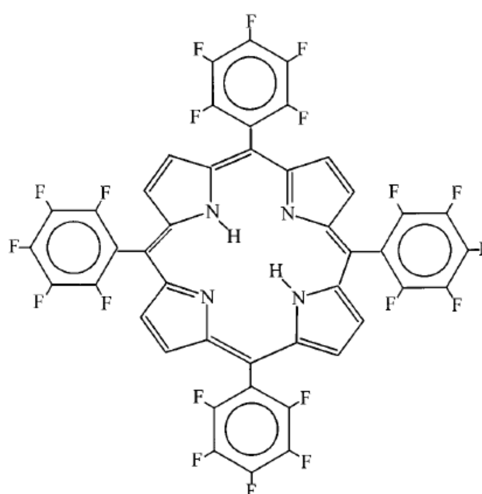


Figure 1. this is the atomic structural configuration of meso-tetrakis(pentaflurophynl)-porphyrin

What do you generally dissolve your sample and matrix in?

Peptides : sample in 0.1% TFA
matrix in 50/50 0.1%TFA/ACN

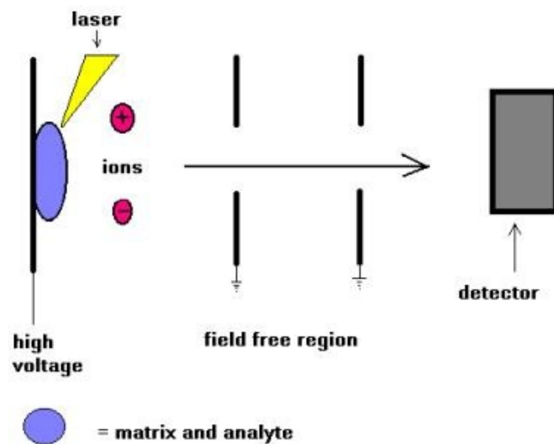
Aids mixing and quickens the drying process

Proteins : sample in 0.1% TFA
Matrix in 70/30 0.1% TFA/ ACN

Higher ACN concentrations precipitate protein

Lipids : sample in chloroform
matrix must also be in miscible organic solvent

There are a number of different small acids that will act as the matrix solutions all of which contain a benzene ring that will absorb the laser energy at the given wavelength. The choose of acid used will be based on the size of the analyte your going to be studying. Peptides are dissolved in an aquiose solution of TFA. For proteins they are also in aqueous solution, to avoid a precipitate of the protein you need to make sure the two mix well. Lipids wont mix in water so they need a chloroform coat to protect it.

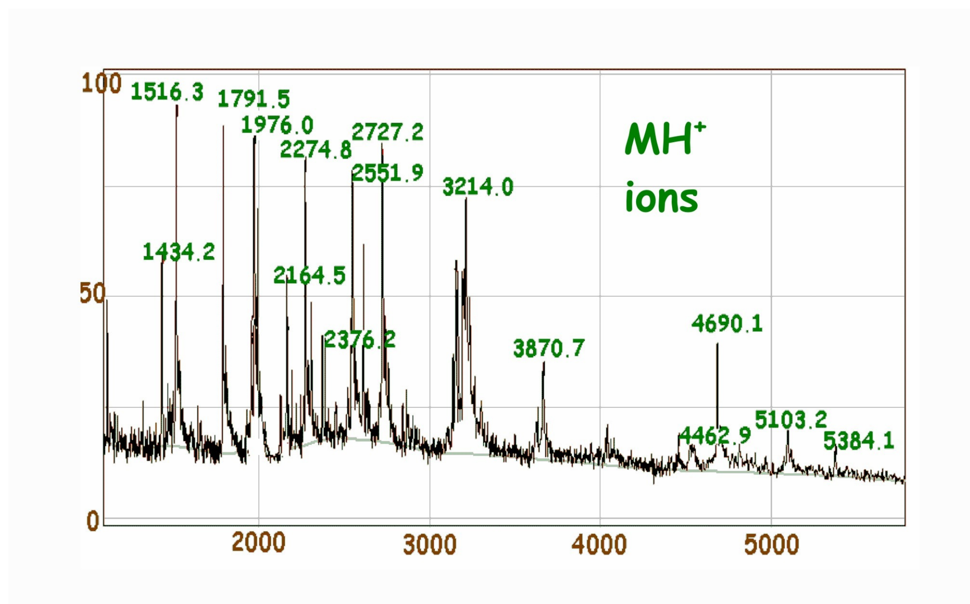


Simplified schematic of MALDI-TOF mass spectrometry (linear mode)

MALDI is also a "soft" ionisation method and so results predominantly in the generation of singly charged molecular-related ions regardless of the molecular mass, hence the spectra are relatively easy to interpret. Fragmentation of the sample ions does not usually occur.

In positive ionisation mode the protonated molecular ions ($M+H^+$) are usually the dominant species, a trace of the doubly charged molecular ion at approximately half the m/z value, and/or a trace of a dimeric species at approximately twice the m/z value. Positive ionisation is used in general for protein and peptide analyses.

In negative ionisation mode the deprotonated molecular ions ($M-H^-$) are usually the most abundant species, accompanied by some salt adducts and possibly traces of dimeric or doubly charged materials. Negative ionisation can be used for the analysis of oligonucleotides and oligosaccharides.



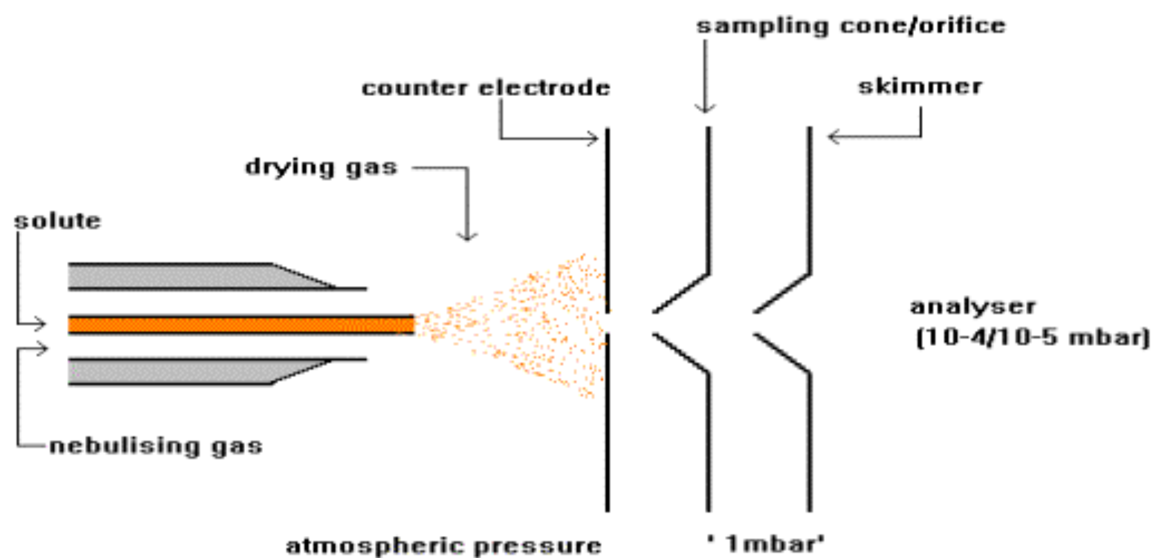
Positive ionisation MALDI m/z spectrum of a peptide mixture using alpha-cyano-4-hydroxycinnamic acid as matrix.

Chapter 2

13. Electrospray

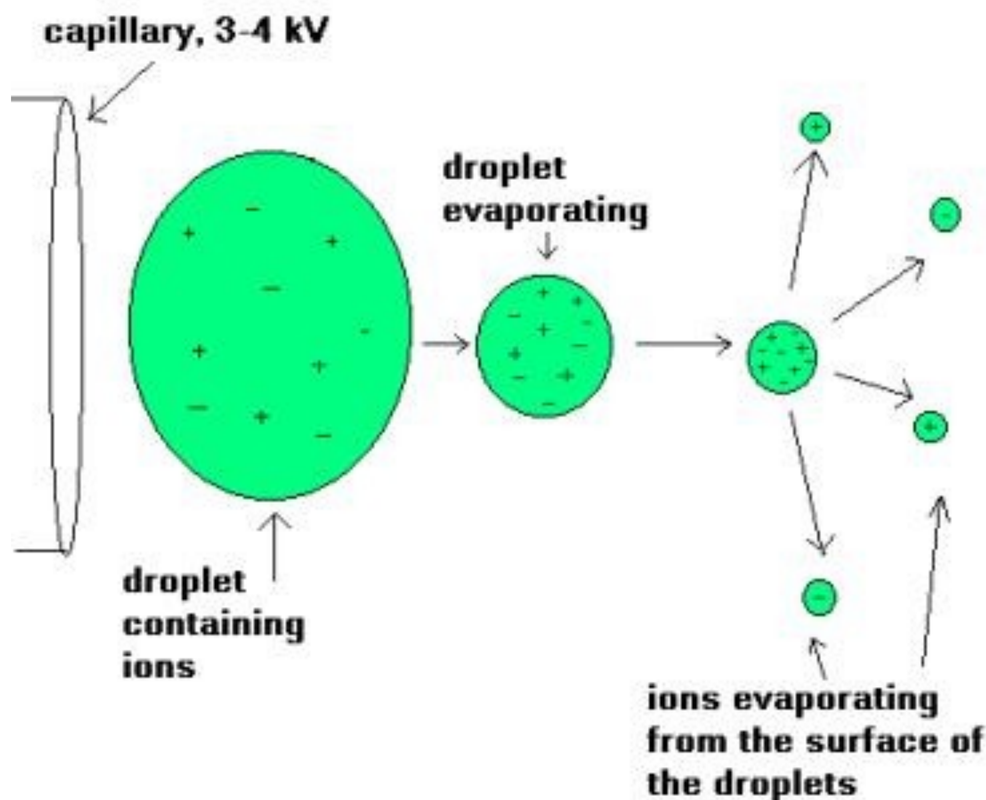
13.1 Electrospray

Electrospray Ionisation (ESI) is one of the Atmospheric Pressure Ionisation (API) techniques and is well-suited to the analysis of polar molecules ranging from less than 100 Da to more than 1,000,000 Da in molecular mass.



Standard electrospray ionisation source (Platform II)

During standard electrospray ionisation, the sample is dissolved in a polar, volatile solvent and pumped through a narrow, stainless steel capillary (75 - 150 micrometers) at a flow rate of between 1 L/min and 1 mL/min. A high voltage of 3 or 4 kV is applied to the tip of the capillary, which is situated within the ionisation source of the mass spectrometer, and as a consequence of this strong electric field, the sample emerging from the tip is dispersed into an aerosol of highly charged droplets, a process that is aided by a co-axially introduced nebulising gas flowing around the outside of the capillary. This gas, usually nitrogen, helps to direct the spray emerging from the capillary tip towards the mass spectrometer. The charged droplets diminish in size by solvent evaporation, assisted by a warm flow of nitrogen known as the drying gas which passes across the front of the ionisation source. Eventually charged sample ions, free from solvent, are released from the droplets, some of which pass through a sampling cone or orifice into an intermediate vacuum region, and from there through a small aperture into the analyser of the mass spectrometer, which is held under high vacuum. The lens voltages are optimised individually for each sample.



The electrospray ionisation process

13.2 Nanospray ionisation
Nanospray ionisation is a low flow rate version of electrospray ionisation. A small volume (1-4 microL) of the sample dissolved in a suitable volatile solvent, at a concentration of 1 - 10 pmol/microL, is transferred into a miniature sample vial. A reasonably high voltage (700 - 2000 V) is applied to the specially manufactured gold-plated vial resulting in sample ionisation and spraying. The flow rate of solute and solvent using this procedure is very low, 30 - 1000 nL/min, and so not only is far less sample consumed than with the standard electrospray ionisation technique, but also a small volume of sample lasts for several minutes, thus enabling multiple experiments to be performed. A common application of this technique is for a protein digest mixture to be analysed to generate a list of molecular masses for the components present, and then each component to be analysed further by tandem mass spectrometric (MS-MS) amino acid sequencing techniques.

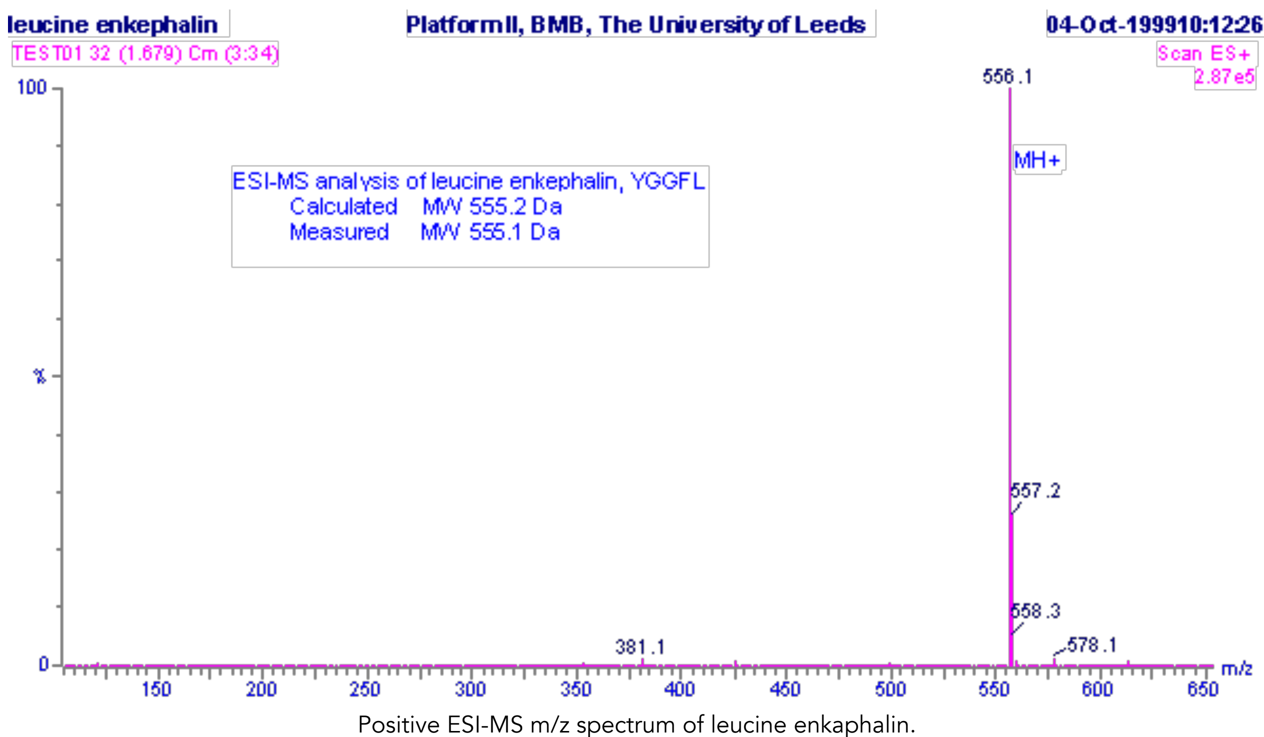
In positive ionisation mode, a trace of formic acid is often added to aid protonation of the sample molecules; in negative ionisation mode a trace of ammonia solution or a volatile amine is added to aid deprotonation of the sample molecules. Proteins and peptides are usually analysed under positive ionisation conditions and saccharides and oligonucleotides under negative ionisation conditions. In all cases, the m/z scale must be calibrated by analysing a standard sample of a similar type to the sample being analysed (e.g. a protein calibrant for a protein sample), and then applying a mass correction.

13.3 Data processing
ESI and nanospray ionisation generate the same type of spectral data for samples, and so the data processing procedures are identical.

In ESI, samples (M) with molecular masses up to 1200 Da give rise to singly charged molecular-related ions, usually protonated molecular ions of the formula $(M+H)^+$ in positive ionisation mode, and deprotonated molecular ions of the formula $(M-H)^-$ in negative ionisation mode.

An example of this type of sample analysis is shown in the m/z spectrum of the pentapeptide leucine enkephalin. The molecular formula for this compound is $C_{28}H_{37}N_5O_7$ and the calculated monoisotopic molecular weight is 555.2692 Da.

The m/z spectrum shows dominant ions at m/z 556.1, which are consistent with the expected protonated molecular ions, $(M+H^+)$. Protonated molecular ions are expected because the sample was analysed under positive ionisation conditions. These m/z ions are singly charged, and so the m/z value is consistent with the molecular mass, as the value of z (number of charges) equals 1. Hence the measured molecular weight is deduced to be 555.1 Da, in good agreement with the theoretical value.



Positive ESI-MS m/z spectrum of leucine enkaphalin.

The m/z spectrum also shows other ions of lower intensity (25 % of the m/z 556.1 ions) at m/z 557.2. These represent the molecule in which one ^{12}C atom has been replaced by a ^{13}C atom, because carbon has a naturally occurring isotope one atomic mass unit (Da) higher. The intensity of these isotopic ions relates to the relative abundance of the naturally occurring isotope multiplied by the total number of carbon atoms in the molecule. Additionally the fact that the ^{13}C ions are one Da higher on the m/z scale than the ^{12}C ions is an indication that z = 1, and hence the sample ions are singly charged. If the sample ions had been doubly charged, then the m/z values would only differ by 0.5 Da as z, the number of charges, would then be equal to 2.

The m/z spectrum also contains ions at m/z 578.1, some 23 Da higher than the expected molecular mass. These can be identified as the sodium adduct ions, $(M+Na)^+$, and are quite common in electrospray ionisation. Instead of the sample molecules being ionised by the addition of a proton H^+, some molecules have been ionised by the addition of a sodium cation Na^+. Other common adduct ions include K^+ (+39) and NH_4^+ (+18) in positive ionisation mode and Cl^- (+35) in negative ionisation mode.

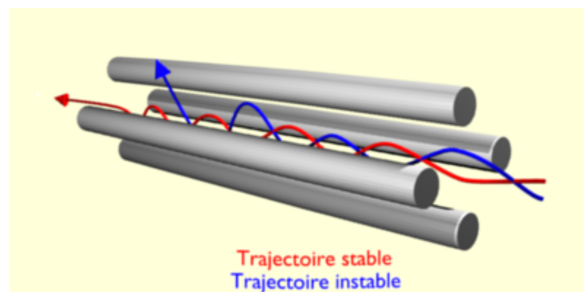
Electrospray ionisation is known as a "soft" ionisation method as the sample is ionised by the addition or removal of a proton, with very little extra energy remaining to cause fragmentation of the sample ions.

Samples (M) with molecular weights greater than 1200 Da give rise to multiply charged molecular-related ions such as $(M+nH)^{n+}$ in positive ionisation mode and $(M-nH)^{n-}$ in negative ionisation mode. Proteins have many suitable sites for protonation as all of the backbone amide nitrogen atoms could be protonated theoretically, as well as certain amino acid side chains such as lysine and arginine which contain primary amine functionalities.

14. Mass Analysers and tandem mass spectrometry

14.1 Quadripoles

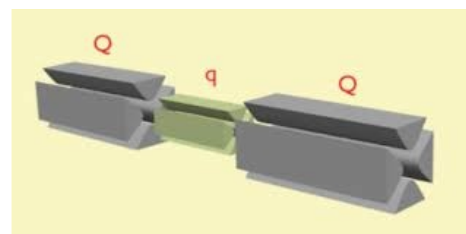
The quadrupole analyser is a device which uses the stability of the trajectories in oscillating electrical fields to separate ions according to m/z ratios. The 2-D or 3-D ion traps are based on the same principle. Quadrupole analyses are made up of 4 rods of circular. The roads must be perfectly parallel.



Quadra polls are pairs of positively and negatively charged circular polls. They are normally configured in a Diamond shape. There have oscillating voltages between them. This is so only certain size particles can travel through them. Alternating currents Will only know how 1m/z value through the tube. The movement of the ion that will travel through the tube Will follow characteristically A sine wave. Ions that do not make it through Will exceed the magnetic fields and collide with the tube. That ions that make it through the tube I called stable trajectories the ones that do not make it through I called instable trajectories. The oscillating currents of the quadrupole can be changed to select different ion masses and or charges. Scanning Quadra polls are used to change the oscillating tyrant from low to high to scan the full range of your sample this can take microseconds to do

A trend in mass analyser development is to combine different analysers in sequence in order to increase versatility and allow multiple experiments to be performed. The five main characteristics for measuring the performance of mass analysers are the mass range limits, that analysis speed, the transmission, the mass accuracy and resolution. The mass range determines the limit of the m/z over which the mass analyser can measure ions. It is expressed in Th or in u for an ion carrying an elementary charge, that is z=1. the analysis speed also called the scan speed, is the rate at which the analyser measures over a particular mass range. It is expressed in mass units per second(u s-1) or in mass units per millisecond. The transmission is the ratio of the number of ions reaching the detector and the number of ions entering the mass analyser. The transmission generally includes ion losses through sections of the mass analyser such as electric lenses before and after that analyser. Mass accuracy indicates the accuracy of the m/z provided by the mass analyser it is the difference that is observed between the theoretical and the measure m/z. it can be expressed in millimass units but is often express in parts per million(PPM). Mass accuracy is largely linked to the stability and the resolution of the analyser. A low resolution instruments cannot provide high accuracy. The precision obtained of the mass of the analysed sample depends also on the determination of the centroid of the peak. Heinemann's accuracy has significant applications such as the determination of elemental composition this is to do with the fragmentation of sample ions to determine structural make up. Lastly is resolution resolution or resolving power is the ability of a mass analyser to yield distinct signals for two ions with small m/z differences. The exact definition of this is to peaks are considered to be resolved if the valley between man is equal to 10% of the weaker peak intensity.

14.2 Triplet quadripole



triple quadrupole mass spectrometer is a tandem mass spectrometer consisting of two quadrupole mass spectrometers in series, with a (non mass-resolving) radio frequency (RF)-only quadrupole between them to act as a cell for collision-induced dissociation. The first (Q_1) and third (Q_3) quadrupoles serve as mass filters. Precursor ions selected in Q1 are dissociated in the collision cell in the presence of an inert gas such as Ar, He, or N_2 collision-induced dissociation. Resulting fragments are passed through to Q_3 where they may be filtered or scanned. This configuration is often abbreviated QqQ, here $Q_1q_2Q_3$.

This alignment of two mass filters with a collision quadrupole between them allows to elucidate the structure of the ionised sample molecules. There are four main methods that can be performed as follows.

• Product ion scan
The first quadrupole Q_1 is set to select an ion of a known mass, which is fragmented in q_2. The third quadrupole Q_3 is then set to scan the entire m/z range, giving information on the sizes of the fragments made. The structure of the original ion can be deduced from the ion fragmentation information. This method is commonly performed to identify transitions used for quantification by tandem MS.

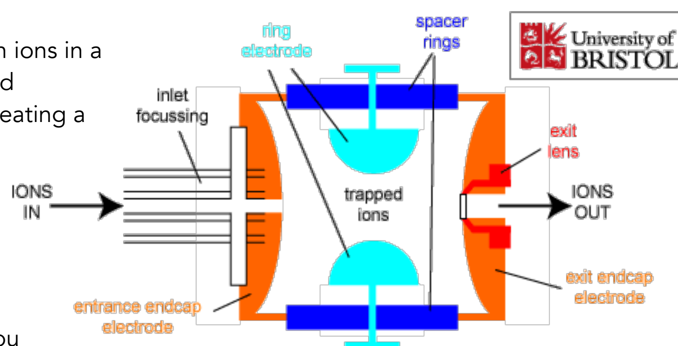
• Precursor ion scan
A certain product ion is selected in Q_3, and the precursor masses are scanned in Q_1. This method is selective for ions having a particular functional group (e.g., a phenyl group) released by the fragmentation in q_2.

• Neutral loss scan
In this method both Q_1 and Q_3 are scanned together, but with a constant mass offset. This allows the selective recognition of all ions which, by fragmentation in q_2, lead to the loss of a given neutral fragment (e.g., H_2O, NH_3). Similar to the precursor ion scan, this method is useful in the selective identification of closely related compounds in a mixture.

• Selected reaction monitoring (SRM) / Multiple reaction monitoring (MRM)
In this method both Q_1 and Q_3 are set to a selected mass, allowing only a distinct fragment ion from a certain precursor ion to be detected. This method results in increased sensitivity. If Q_1 and/or Q_3 is set to more than a single mass, this configuration is called multiple reaction monitoring.[2]

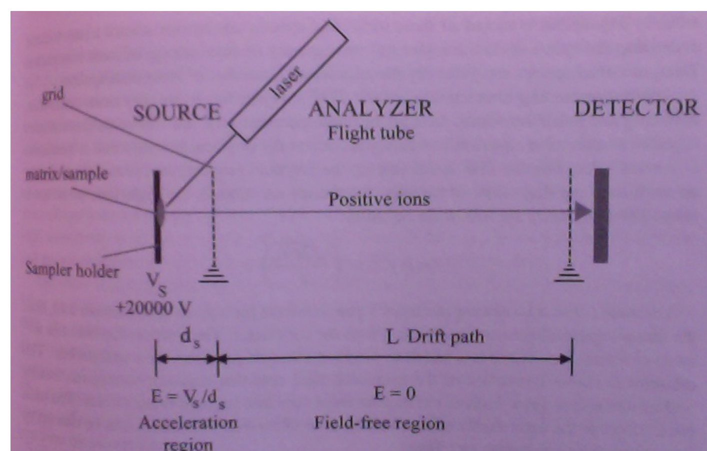
14.3 Ion Traps

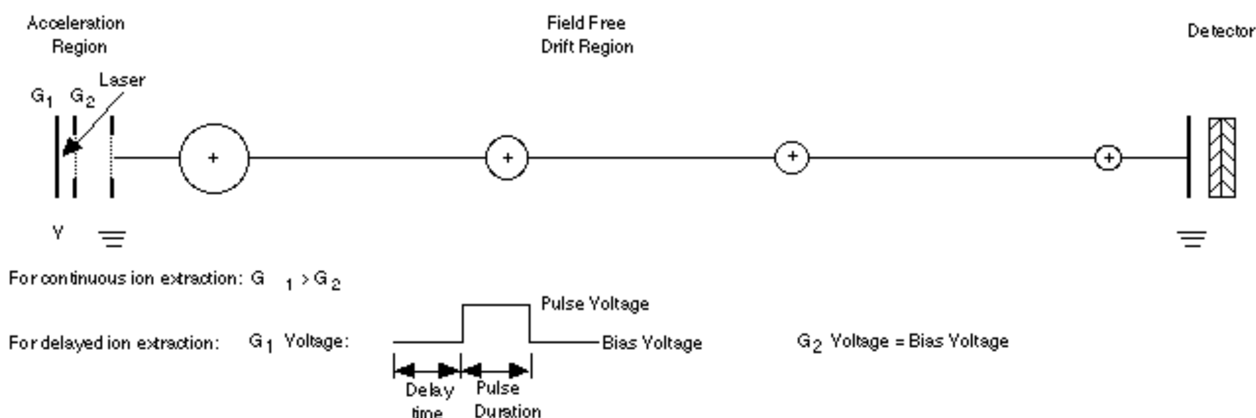
An ion trap is like A cross section of a quadrupole used to contain ions in a box. The ions will oscillate in a continuous pattern surrounded and maintain by oscillating electrical fields. You capture the ions by creating a stable electrical field. Disrupting this field you can essentially disrupt the circular pattern or figure of eight pattern that the ions are following and fragment the ions. Ions can be trapped and classified into types of 3-D ion trap or the 2-D ion trap. Historically the first traps were 3-D I am trapped stop there were made up of circular electrodes, with two caps on the bottom and two caps on the top. The general principle all I am traps is that you can do multiple MS/MS. quadrupole's separate a precursor ions and then separate the fragment ions. With an ion trap you can isolate an ion and fragment it. You can then take bees fragments and analyse them individually by further fragmentation.



14.4 Time of flight tubes

Time of flight tubes (TOF) is a cylindrical tube in which an ion's mass-to-charge ratio is determined via a time measurement. Ions are accelerated by an electric field of known strength. All ions of varying size are attracted to this electric field and the will have same attraction force dependent on there charge (friction will not be a factor) but will have different accelerations dependent on there mass. This acceleration results in an ion having the same kinetic energy as any other ion that has the same charge however there velocities when they have reached there top speed depends on the mass-to-charge ratio. The time that it subsequently takes for the particle to reach a detector at a known distance is measured. This time will depend on the mass-to-charge ratio of the particle (heavier particles reach lower speeds). From this time and the known experimental parameters one can find the mass-to-charge ratio of the ion.





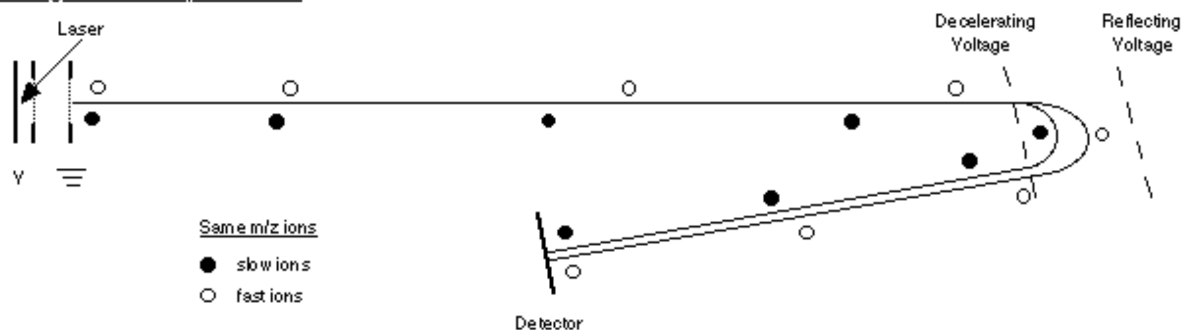


Figure 1. Basic components of a linear (upper) and reflecting (lower) TOF mass spectrometer.

There are two types of time of flight tube. The first is a linear Time of flight tube where the time is measured for the particles to reach one end of the tube to the other. The other is a reflecting time of flight tube where the particles will travel into an electrical field where they are reflected back on their path at an angle. Heavier particles or not deflected as easily as small particles and will either keep travelling or will deflect at lower angles. Lighter particles generally below 5000Da Will deflect and be measured by the detector. The advantage here in reflecting timer flight tubes is that the longer they travel the more separated the ions will become. This generally corresponds to greater resolution and better mass accuracy.

Time-of-flight mass spectrometers are based on a simple mass separation principle. Consider ionised species starting from the same position at the same time, being accelerated by means of a constant homogeneous electrostatic field. Their velocities are unambiguously related to their mass-to-charge ratio and times of arrival at a detector directly indicate their masses (the mass isn't actually measured but the accurate flight time corresponds to an accurate mass:

$$t = \left(\frac{2md}{eE}\right)^{\frac{1}{2}} + L\left(\frac{m}{2eV_0}\right)^{\frac{1}{2}}$$

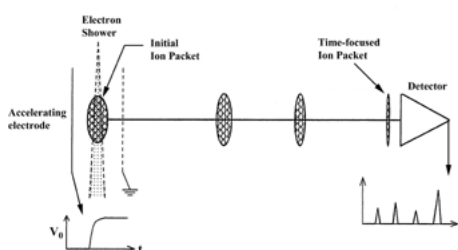
where m = mass of particle, e = electronic charge, E = electrostatic field applied in source, d = length of accelerating region, L = length of field-free region, and Vo = accelerating potential.

The principle of time-of-flight has been known since Thomson carried out his experiments on ionised particles. The first proposal, however, for a mass spectrometer based on the time-of-flight principle was made by Stephens.

The time-of-flight instrument possesses a number of extraordinary advantages over most other types of mass analyser:

* Theoretically unlimited mass range.
* Ideal where ionisation is pulsed or spatially confined.
* Complete mass spectrum for each ionisation event.
* High transmission.
* No need for scanning the ion beam (the Fellgett advantage).
* Spectra can be obtained for extremely small sample amounts (<10-18 mole in the most modern instruments).
* Relatively low cost.

Wiley and McLaren observed that ions of a particular mass-to-charge ratio would reach the detector with a spread in arrival times, due to the effects of uncertainty in the time of ion formation, location in the extraction field and initial kinetic energy, resulting in reduced resolution. Wiley and McLaren devised an instrument, incorporating a pulsed two-grid ion source, to compensate for temporal, spatial and initial kinetic energy distributions. The basic geometry of the Wiley-McLaren design is shown in the figure below.



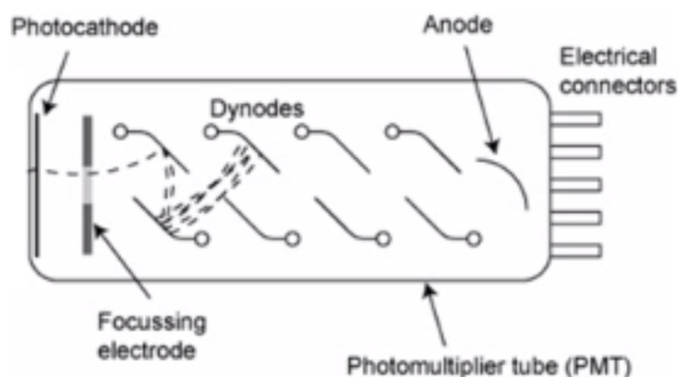
14.5 QTOF

Quadruple time of flight tubes. This is an MS/MS technique that looks at increasing the sensitivity. Using a time of flight tube allows you to separate masses based on the flight times. Using a reflecting TOF tube increases the separation of the particles and thus increases the resolution of the peaks. Using a reflecting tube decreases the sensitivity though. The longer the flight time the less accurate it will be. Introducing a quadruple in stead of another time of flight tube means that the particular ions can be separated and fragmented but the accuracy about what ions are selected are increased.

---

15. Electron Multiplyers

Electron multiplyer tubes work essentially by having a cathode that an electron will smash into, this will then send an electron through a focusing slit which will hit a dynode. A dynode is a material much like a wall covered in sticky tape and a bunch of ball barings. When one ball barring hits the wall at high speed then two or three others may fall off. The same principle applies to a dynode with an electron. An electron will hit it from the cathode and create a cascade of electrons between a number of dynodes. This tube will eventually end in an Anode which attracts all the electrons and can detechet the increases signal. The Increased signal can then be related to the number of initial electrons that hit the cathode. this essentuly develops a charge that is detectable by a computer.



Chapter 3

This chapter includes the detailing of the pre MS separation methods including:

- Gas chromatography
- High pressure liquid chromatography
    - HPLC retention
    - Capacity Factors
    - Selectivity
    - Efficiency
    - Resolution
    - Types
        - Reverse Phase
        - Normal Phase
    - Elution Types
    - HPLC miniaturisation

---

16. Seperation Methods

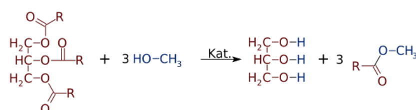
You normally add a separation instrument onto the inlet of the MS machine. This could be something like a HPLC machine or GC. You make the experiment more convoluted with the addition of these machines as you now adding a module phase into an analysing machine. With electrospray you can help purify your sample. The main reason you would want to do this is because things like electrosprays only have the ability to ionise a certain amount of sample at one time. If you had two compounds in a sample an one is 100 times more concentrated that the other you wouldn't see the other compound in the detection phase. individually separating the molecules first allows the individual identification of the molecules. An example of a real wold cenario is the analysis of things like proteins and DNA. The separation methods for these samples require the use of salts and buffers which can affect the ionisation and detection of the correct compounds. So by separation first you have a clear path to the analysis you want.

16.1 Volatility

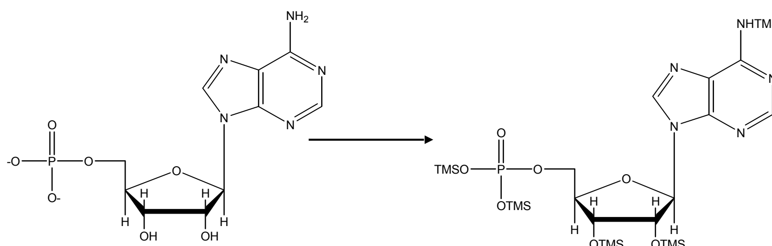
Adding a material to a seperation apperatus before it reaches the MS requires that the compound be in gas phase. This is normally done by injecting the sample into a heated chamber which vaporises the sample and is then added to the GS or HPLQ with an inert gas like argon or xenon. Some compounds like lipids are not in a gas phase and are thus not volatile and wont vaporise or wont stay vaporised. This can be fixed by adding compounds to create derivatives of the analyse that are more volatile.

Examples.

- With fatty acids you can add methyl esters



- With other compounds you can add TMS (trimethylsilane). You add TMS to replace all of the hydrogen atoms on the compound.



16.2 Columns

A column in a GS or HPLC apparatus is used as the stationary phase. This means that the column will be cited with beads like silicon beads which will have a certain property about them. They can either have different sized cracks in them that can catch certain sized molecules, they can be attached to charged particles that will have different affinities for different molecules and more. We will learn more about these later. For now you need to know that the mobile phase is the column and it traps certain molecules. There will be a lumen to the column kind of like a blood vessel which will allow the mobile phase to pass through. You will inject your sample into the mobile phase and wether it has an interaction with that stationary phase or not. You change the nature of the mobile phase to reduce the affinity for the sample to interact with the stationary phase. specific control of this change to the mobile phase can allow you to release the compounds one at a time.

16.3 Capacity Factors

The first thing you will want to understand is wether your compound interacts with the compound at all. You can do this by determining the capacity factor K'

Major parameters,

- ***V*** is retention volume, *depends on the column type, size, and the instrument parameters – also Rt*
- ***V***^{R} is dead volume, *volume of the liquid phase inside the column*
- ***k***^{9} is retention factor (capacity factor),

$$k' = \frac{V_R - V_o}{V_o}$$

Dead volume = The volume of liquid that sample has to pass through in order to pass through the column. Retension volume is the amount of liquid that has to pass through the column to elute the sample.

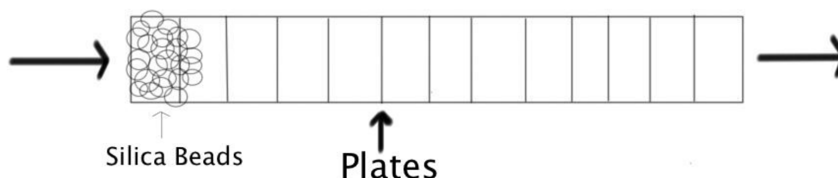
From the equation you can find the retention volume and the longer a compound sticks the column the higher the capacity factor will be. This type of equation takes out the factor of the column diameter by normalising to the dead volume.

Usually you have more than one compound and if so you will need a selectivity higher than 1.2. Selectivity can be found from the following equation.

$$\alpha = \frac{k_2^{'}}{k_1^{'}};$$

16.4 HETP (hight equivalence of theoretical plate)

To find the selectivity of the column you can divide the capacity factors by one another as seen in the equation above. You could also find the efficiency of the separation of the compounds which is determined by the number of theoretical plates in a column.



The theoretical plates don't exist but you can determine the theoretical number of them by the squation below

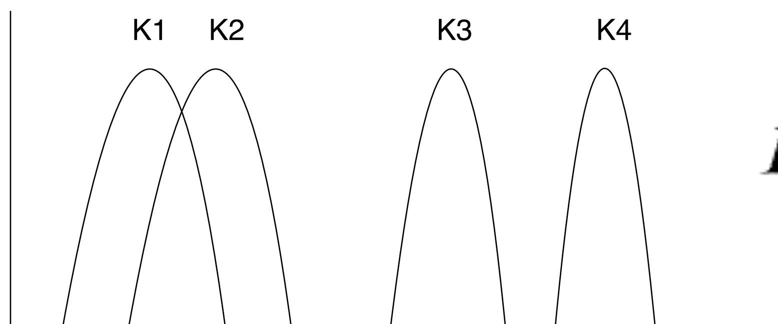
$$N = 16\left(\frac{V_R}{w_b}\right)^2$$

This just shows the retention volume decided by the width of the peak (Wb). So efficiency defined by the number of plates and the width of the peak. Thus if the peak is wide then the efficiency ist very good and the higher the retention volume is then the better it is. The more retention volume you can get over a smaller peak width the more effient it is. This can also relate to the sensitivity as the better the efficiency (the smaller the width) the higher the peak must be to contain the same number of area units under the peak. higher peak means more sensitive.

Each compound has a theoretical plate number (N). the theoretical height of the peak in the spectrum can be determined by taking into account the length of the column. Reducing HEPT means that the column isn't as well packed.

$$HETP = \frac{L}{N}$$

16.5 Resolution



$$R = \frac{\Delta t}{\frac{1}{2}(w_1 + w_2)}$$

As you can see from the two peaks the selectivities are the same. This is because the peak apexes are at the same height. However the separations of the retention times between K3 and K4 are different to K1 and K2. Clearly there is better separation between the the K3 and K4 so we say the resolution is better and this is found by the equation to the right of the diagram.

---

17. Reverse Phase Principle

17.1 Reversed-phase separation principle

This includes any chromatographic method that uses a hydrophobic stationary phase. RPC refers to liquid (rather than gas) chromatography.

In the 1970s, most liquid chromatography was performed using a column containing unmodified silica resins. This method is now called "normal phase chromatography". In normal phase chromatography, the stationary phase is hydrophilic and therefore has a strong affinity for hydrophilic molecules in the mobile phase. Thus, the hydrophilic molecules in the mobile phase tend to bind (or "adsorb") to the column, while the hydrophobic molecules pass through the column and are eluted first. In normal phase chromatography, hydrophilic molecules can be eluted from

the column by **increasing the polarity** of the solution in the mobile phase. The introduction of a technique using alkyl chains covalently bonded to the solid support created a hydrophobic stationary phase, which has a stronger affinity for hydrophobic compounds. The use of a hydro*phobic* stationary phase can be considered the opposite, or "reverse", of normal phase chromatography - hence the term "reversed-phase principle". Reversed-phase chromatography employs a **polar** (aqueous) mobile phase. As a result, hydrophobic molecules in the polar mobile phase tend to adsorb to the hydrophobic stationary phase, and hydro*philic* molecules in the mobile phase will pass through the column and are eluted first. Hydrophobic molecules can be eluted from the column by **decreasing the polarity** of the mobile phase using an organic (non-polar) solvent, which reduces hydrophobic interactions. The more hydrophobic the molecule, the more strongly it will bind to the stationary phase, and the higher the concentration of organic solvent that will be required to elute the molecule. It can be used for the separation of a wide variety of molecules. It is not typically used for separation of proteins, because the organic solvents used in RPC can denature many proteins. For this reason, normal phase chromatography is more commonly used for separation of proteins. Today, RPC is a frequently used analytical technique. There are a variety of stationary phases available for use in RPC, allowing great flexibility in the development of separation methods.

An example of the gradient elution of the RPC method can be seen in the table below. By increasing the methanol concentration in the mobile phase you elute more coumpounds. At 20% methanol A,B,C and D

Mobile phase	Retained
20% meoh	A,B,C,D
30% meoh	B,C,D
40% meoh	C,D
50% meoh	D
60% meoh	

are all bound to the sataionary phase column. By increasing the MEOH concentration to 30% A is knocked off the column and is eluted. Increasing the MEOH concentration again elutes morte and at 60% none of the compounds will stick to the column any more.

Reverse phase is most commonly used but you can also use:

1. Normal phase partitioning
2. Absorption partitioning
3. Ion exchange
4. Gel filtration

Reverse phase is most commonly used in HPLC though as salts and buffers cause problems for the MS machines. Transporting sample analytes in a reverse phase system allows you to eliminate the salts and buffers.

17.2 HILIC

HILIC stands for hydrophilic interaction cromatography. HILIC is used for the separation of polar and hydrophilic compounds. Stationary phases are polar similar to normal phase chromatography. Typical HILIC phases are silica or polymer particles carrying polar functional groups. Mobile phases are aqueous buffers with organic modifies applied (usually 95% acetyl nitrile) in isocratic or gradient mode. The gradient mode would work by increasing the water concentration from 5% to encourage the release of compounds from the stationary phase.

Aquaeouse contents of the mobile phase in HILIC creates a water rich layer on the surface of the stationary phase. This allows portioning of elutes between the more organic mobile phase and the aqueous layer. The number of olar groups and the conformation and solubility of the sample in a the mobile phase determines the elution order. Compared to Reverse phase the elution order in HILIC is inversed for most compounds. HILIC is ideally suited for mass spectrometric analysis of water soluble polar compounds. Because the high organic content in the mobile phase increases MS detection sensitivity.

The advantage of HILIC is that your now eluting in a 95% acytlnitrile solution which is much easier to evaporate in an electrospray apparatus. Acytlnitrile is a lot more volatile which means is significantly easier to get to the ionisation source than water.

---

18. HPLC miniaturisation

Traditionally HPLC uses columns of diameter 4.6mm and run at 1ml/min. In Nano-HPLC ou shrink this to a diameter of 75micrometers and run at 200nl/min. The reason for reducing the size of the HPLC columns is because the elctrospray ionisation at the MS inlet is a concentration sensitive detector. Reducing the amount of mobile phase reaching the electrospray incases the ionisation effiencey of the elctrospray example.

5nmoles injected into the column

elutes in 1min peak from the HPLC

1ml/min : 5nmoles/1ml = 5uM concentration

200nl/min : 5nmoles/200nl = 25nM concentration

An electrospray ion sourse is more effective at ionising a larger concentration in a small volume because there is a discreet desolation effect where the analyses dissolve in the mobile phase and wont be ionised.

The changes with miniaturising an HPLC is that your larger compounds like proteins and larger peptides will now be detected as multiply charged ions and you wont be able to directly inject your sample into the column as it will overwhelm the system. Instead you use a trapcolum with an autosampler which isolates your analtyes in a trap column which is bound to the HPLC later.

Lecture 4

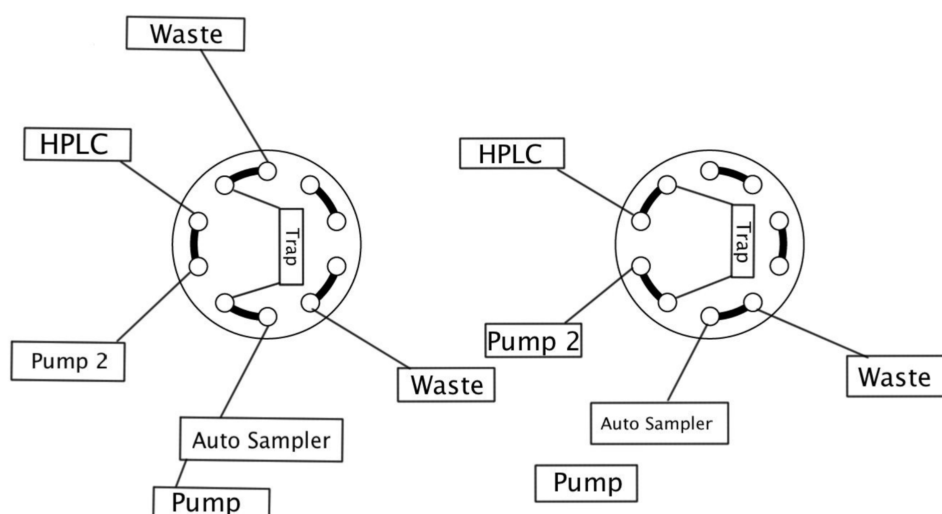
Chapter 4

This chapter includes the detailing of:

- The ten port valve system
- The six detection methods in HPLC/MS (advantages and disadvantages)
  - Full Scan (FS)
  - Single ion Monitoring (SIM)
  - Single Reaction Monitoring (SRM)
  - Constant Neutral Loss (CNL)
  - Precursor ion discovery (PID)
  - Data dependent acquisition (DDA)

---

19. Ten port valves



In the left hand trap system. The ten port valve system is receiving the analyte of interest from the pump and thats going into an auto sampler. The auto sampler is a commonly used tool today because it can be very accurate and reproducible. The rate and quantity of the analyte can be controlled accurately and can be identically done many times over. This autosampler is loading the analyse into a trap which is a small column. The trap is a C18 column with a small internal volume that is pours. Most the compounds of interest can be concentrated onto the C18 column trap. Pump 2 is loading liquid at 0.2ul/min into the HPLC system but isn't loading any of the analyse into the HPLC. The 10 port trap system is then switched and Pump 2 is now moving the mobile phase through the trap into the HPLC system. The HPLC is a separation column and has a low flow rate. In the interest of saving time in the separation of the analytes, the trap concentrates the analyte sample into its dead volume so that they can be eluted into the HPLC in one go. You have affectively taken a 5ml volume and concentrated it into a 3nl volume on the trap which means the analyse can now more effectively be measured with the low flow rate. This reduces the time taken for the trapment of the sample onto the HPLC column from 25 minutes to half a minute.

20. Detection methods

There are six fundamental detection mechanism for detecting ions. Each one offers different advantages and disadvantages as they offer slightly different methods. Depending on what you know about your ions or what your looking for you would choose one of these methods. From the list below the last four use tandem MS where you will select a precursor ion and collide it with an inert gas to fragment it. The first two use a full scan and a single MS. You can also do multiple reaction monitoring with a few of these methods. multiple reaction monitoring (MRM) is where you tell the detectors that you will be looking for different ions at different times. therefor it can adjust its parameters for the different ions expected at different elution times from the HPLC.

- Full scan (FS)
- Single Ion Monitoring (SIM)
- Single Reaction Monitoring (SRM)
- Constant Neutral Loss (CNL)
- Precursor Ion Discovery (PID)
- Data Dependent Acquisition (DDA)

20.1 Full Scan

Full scan analysis is a single reaction monitoring. you measure an entire spectrum and detect all ions present wether your interested in them or not. this would be good for a lot of comparative analysis to ask if too samples had the same ion characteristics or to ask about the specific composition of the sample. Example, did it have testosterone in it, did it have magnesium, did it have this specific protein. We dont need to know anything about the sample or have any specific interest in the particulars of the sample. from a full scan you get a **total ion chromatogram (TIC)** which measures the abundance of all the ions in the sample and displays then in relation to each other. This will however not show any abundant peaks. you can also get an extracted ion chromatogram from a full scan which looks for selected ions in your scan to determine the presence of particular ions.

Advantages

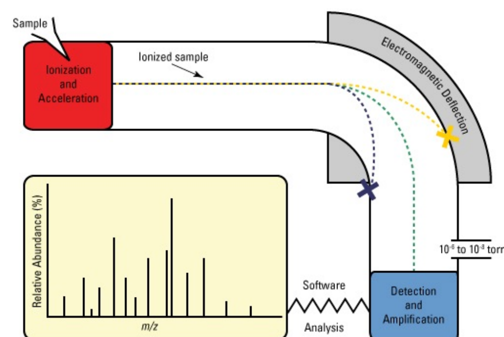
No data is lost and the mass of your sample need not be known and can be extracted through an ion chromatogram after analysis.

Disadvantages

Less sensitive - When you are studying all the m/z values you waste time on the unimportant ones and rant as specific to the m/z you may want.

20.2 Single ion monitoring

This is like a full scan analysis accept only one ion is sent to the detector. All the ions enter the mass spectrometer but there is a magnetic bending parameter that only one ion will be able to meet. this is much like the original MS pictures. This is like extraction ion chromatography but is has an increased sensitivity as it is looking for a specific ion only. not a specific ion in a cluster of ions. A mass separation filter like a quadruple is responsible for the detection of only one ion. You can also get the machine to switch between different time points with a shift in about 50milliseconds. This is useful for compounds that will elute from the HPLC at the same time. Switching like this can prove the existence of both but reduces the sensitivity.



Advantages

There is an increase sensitivity in the ion you are monitoring and the peak width will not be as large.

Disadvantages

The machine will detect any thing at that specific m/z value regardless of wether its the analyte your looking for or not.

20.3 single reaction monitoring
This now incorporates the use of tandem MS. There are three two quadruples and one collision cell. The first quadruple isolates a precursor ion which will enter Q2. Q2 acts as the collision cell by bombarding the precursor ion with an inert gas. Then all the fragment ions enter the Q3 cell where they are separated and a specific ion is isolated. You actively select the precursor ion like you would do with Single ion monitoring (SIM). By fragmenting it and selecting a different sized fragment in Q3 to the precursor ion you eliminate the background noise.

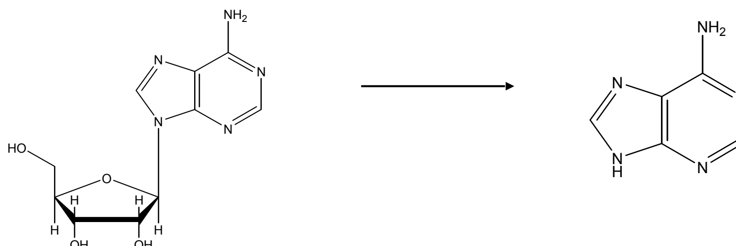
Advantage

Increased selectivity and also sensitivity by reducing the background noise. This is a method commonly used in pharmaceuticals.

Disadvantages

Only the compound of interest can be monitored.

20.4 Constant neutral loss (CNL)
here you have a tandem MS again but you don't select an ion you let the MS select the biggest ion r second biggest and let it fragment it itself. Ask it was there a specific loss when the compound was fragmented if so they you may have identified what your looking for. then only if it obeys your desired m/z do you keep the data. If it dosent match the expected loss in mass then the data isn't valid and is discarded. An example of the type of loss we would look for would be like with and adenosine nucleotide.



Advantages

Increase specificity of analysis that can also look for unexpected compounds in a given class.

Disadvantage

Less sensitive than SRM as it requires a scan of all the ions. It also relies on all the compounds of interest acting in the same manner when it comes to fragmentation. There will be preferred fragmentations of a impound based on it orbital chemistry but they have the potential to fragment in different ways resulting in skewed results.

Lecture 4

20.5 Precursor ion discovery
Here you ask the machine to look for a specific Q3 fragment to determine the precursor ion.
you do tandem again and you look for a specific m/z value in Q3 being specific about which ion was produced. you don't need to know the precursor ion of the fragment you are looking for. This would mean looking for specific fragments of a compound.

CNL and PID use tandem MS but don't look for a specific compound (like SRM) Instead they "recognise" characteristic fragmentations that identify specific types of compounds They therefore balance the specificity and sensitivity of SRM with the analysis of more than one compound and even unknown modifications of known species.

20.6 Data Dependent Acquisition
This method is most commonly used in proteomics You take a protein, Digest it and separate the peptides with HPLC. It essentially tires to determine as accurately as possible by gathering as much information as it can, the peptide that was initially digested. This is achieved by a mass of fragmentation data where you don't make any assumptions abut what is or isn't important collecting data in an unbiased manner. This would then be compared against data bases to determine the peptide sequence and thus the parent peptide.
You would run this on an HPLC in a 60 minute frame and split the time up into 1 second portions. With these 1 second portions you do four things.

1. It dose a full scan mas spectrum
2. detects the three most abundant ions.
3. It then fragments those three ions.
4. repeats this every second.

20.7 Analysis
It dose this 180 times a minute and 11,000 times an hour. Giving you a lot of fragmentation data. This generates a lot of data which would be ineffinet to sift through by humans. So you then go through bioinformatics software to determine the relevance of the fragmentations you derived. You have six methods available to you but you will have to determine when to use which ones and what information you want to relveale from it. You must first ask two questions.

1. What do you want to know
2. How much do you already know

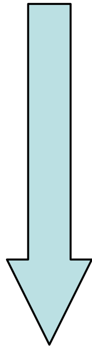
You would determine this from the type of data you have. either quantitative or qualitative. Quantitative is trying to assay or determine the concentration of a compound in a sample. Qualitative looks at determining whats in the sample.

If you want to do quantitative analysis which should you use.
- Full scan good if dont know what you want to identify (unknown sample)
- SIM / SRM these are optimal for quantifying already known compounds (better if known).
- CNL / PID these two methods can help quantify a compound but they can also identify specific structural configurations of a compound. (they can quantify adenosine but also tell you if its methyl adenosine or not in the middle).
- DDA – may waste too much time because it collects so much data that its never going to give you any quantitative data of any value

If you want to do qualitative analysis what should you use.
- SIM / SRM – too specific
- CNL / PID – good for classes of compounds
- FS / DDA – will study everything DDA in more depth but may Sacrifice molecule coverage

Lecture 4

I know what it is



I don't know what it

SRM

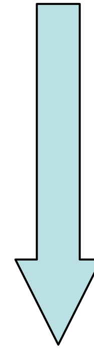
SIM

CNL / PID

FS

DDA

Quantitation



Qualitative analysis

Lecture 5

Chapter 5

This chapter details:

- Proteomic analysis
- MALDI analysis for proteomics.

---

21. Proteomics introduction

Usually in protein (proteomics) and peptide (peptidomics) identification you would use a MALDI-TOF setup. Usually this type of analysis is soft compared to other form of analysis and will only ionia a peptide or protein once. Even larger peptides or proteins will only get a single ionisation. Three uses of this MALDI TOF analysis are:

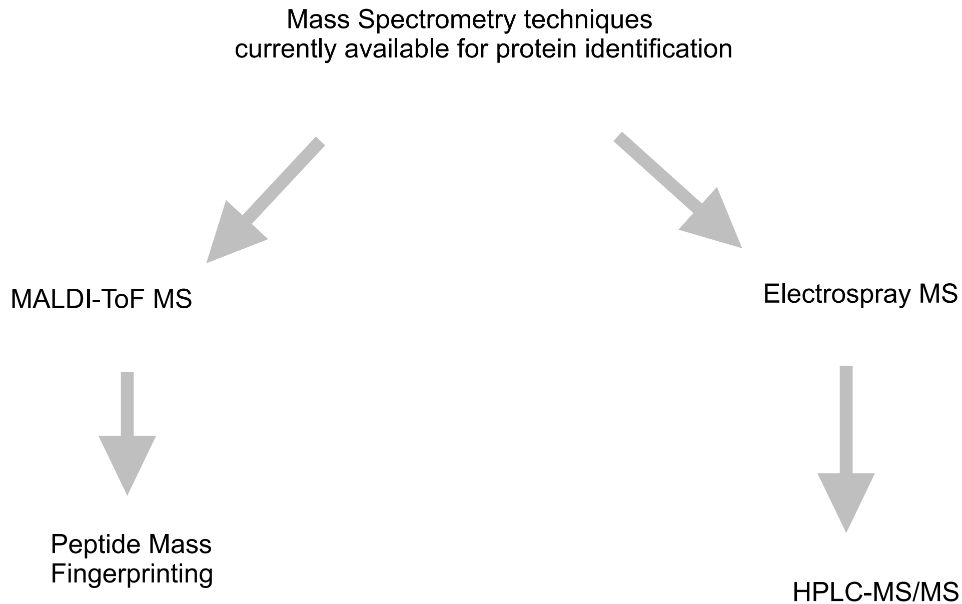
- Micro-organisim Identification
- Disease biomarker discovery
- MALDI imaging

Micro-organisms can be identified with this technique by scraping a colony of a petri dish, ad it to a MALDI plate mixed with a matrix and fundamentally what happens is an dissociation of proteins from the cells that can be studied to produce profiles for each species. For example Ecoli will display more perm ease proteins than S auras. You can actually buy a libraries of a vast number of species and varieties among species like the many types of S aurius. This will be helpful when trying to identify different strains resistant to certain treatments. Different resistant strains will have subtly different profiles and will which can be identified against a database library. This is another form of bioinformatics. The problem with this is the mono isotopic ion resolution where you get different masses of the same ion for its different isotopes. You generally get average resolutions in this type of analysis which means you assume many proteins are what they appear to be but could have resistance modifications that are no different or marginally different in mass and therefor undetectable. Its more of a tool to help species identification than identification a treatment resistant strains.

Disease biomarker discovery helps to identify different proteins or peptides that may be present in the blood serum, blood, sputum or other fluids. You may have a biomarker that is different in a cancer patient compared to a control patient. You may not necessarily know the identification of that protein or biomarker although you may know its presence reflects a disease. This technique is heavily applied to cancer and with other psychiatric disorders like schizophrenia or depression. The problem with looking at blood samples is that it contains a lot of serum albumin which is a 66Kda protein. This will have a large appearance in any scan so the removal of it can help the detection of other biomarker proteins in the blood sample. You separate the serum by clotting the haemoglobin and spinning it to separate the clot. You would then use column affinity to separate the serum. The main barrier is that albumin may attach to some proteins in the blood and thus they will also be removed wight the serum albumin. There are alternative ways of getting rid of it. You can precipitate it with adding different levels of organic solvent. Again its not 100% affective and will lose some biomarkers. The solvent may also degrade the proteins and cause oss of functionality. An alternative strategy is to use auto filtration where you add a gradient with a filter set just below 66Kda. When spun on a centrifuge the smaller proteins would pass through but the larger ones won't. This is similar to the kidneys function this docent degrade the proteins but dose limit the the molecular weight of the proteins to be analysed.
Alternatively instead of separating the proteins from a blood sample the urine. The kidneys are an effective filtration system that can filter out proteins into the urine. The main advantage of taking urine is being able to measure a stress test. a serum or urine sample will not only contain the biomarkers of the potential diseases but will also contain the composition of the person which there currently in. Urine is generally more reliable as a stress towards needles can change the blood composition. Urine dose also have an exported protein called urea which is a problem in the same way serum albumin is for blood samples.
Urine will contain things called exoduses though which are proteins derived from the endometrial cells but they don't have this problem of a varied abundance range. You separate them from sucrose gradients.

22. MALDI protein identification

There are two techniques used for protein identification using mass spectrometry.



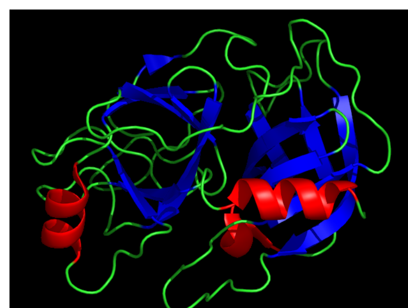
Protein mass fingerprinting is the process of digesting a protein down into peptide sequences of amino acids. These are the sample proteins that would enter a mass spectrometer. These peptide sequences will create a mass list which will be characteristic like a fingerprint to the parent protein it came from. You could separate the proteins before they enter the mass spectrometer not using an HPLC system but a simple SDS PAGE experiment of a 2D PAGE where the migration is based on isoelectric points in a medium based on pH. The key is to digest your protein in a controlled manner so that the protein is characteristically digested at certain points that wont be similar to the points of other proteins. As example of a enzymatic digestion is the use of trypsin whig is universally used and digests it a argentine (R) and lysine (K) point in the peptide chain. You usually treat the peptide with two things first. DDT and iodoacetamide, DDT brakes disulphide bonds in a protein adding a proton to each sulphurs stopping them reform a bridge. This brakes the three devotional structure to create a linear protein. The usually protocol for digestion time is placing them in trypsin at 37 degrees for 16hours and 24 minutes. This is done because a 9 to 5 job normally has a sixteen hour brake from 5 one night to 9 the next morning. the digestion time can be decreased by increasing the temperature of digestion to 55 degrees or place it in an immobilised trypsin column. Once you have digested your proteins with trypsin you can compare the mass list to a database of known peptide masses and of those you can determine which of those looks most like my experimental data .

Chapter 6

This chapter will cover:

- Trypsin
- The calibration theory
- ABC and XYZ fragmentation

23. Trypsin



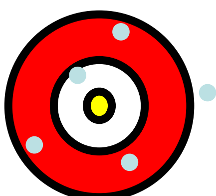
Trypsin is a serine protease from the PA clan superfamily, found in the digestive system of many vertebrates, where it hydrolyses proteins. Trypsin is produced in the pancreas as the inactive proenzyme trypsinogen. Trypsin cleaves peptide chains mainly at the carboxyl side of the amino acids lysine or arginine, except when either is followed by proline. It is used for numerous biotechnological processes. The process is commonly referred to as trypsin proteolysis or trypsinisation, and proteins that have been digested/treated with trypsin are said to have been trypsinized.

24. Calibration theory

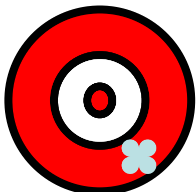
There are two things to consider when talking about calibration. Precision and accuracy.

Precision is a measure of repeatability, i.e. the degree of agreement between individual measurements of a set of measurements, all of the same quantity.

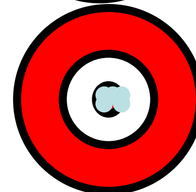
Accuracy is a measure of reliability, and is the difference between the True Value of a measured quantity and the Most Probable Value which has been derived from a series of measures. The True Value is, of course, never known. You increase the accuracy of your data by more accurately determining more decimal places. As your matching experimental data against theoretical data the more accurate you can be the less likely you are to make a mistake. The more accurate you are the fewer matches you will get against your experimental data.



None of the darts are close to the true value (bull's eye) : the measurements are not accurate. Also, since the darts are not very close to each other, the set of measurements is not precise either.



Since all of the measurements are close together, they are precise, but since they are not close to the true value, they are not accurate



The measurements are all close to the true value, so they are accurate. Also, the measurements are all close to each other, so they are precise

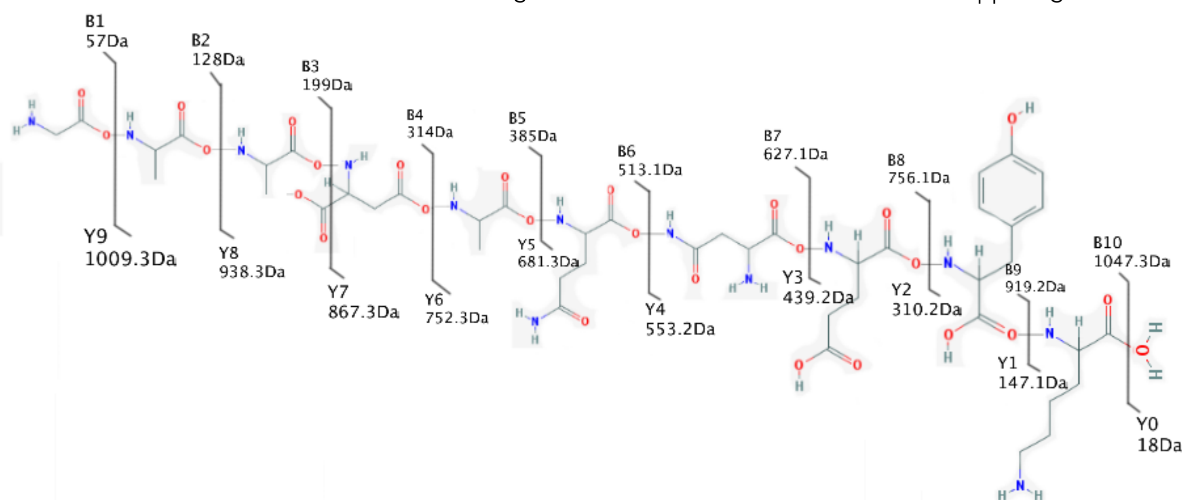
Lecture 6

In databases like the swiss institute of bioinformatics the accuracy and perdition of your data will be displayed in a mowse score. The higher the score the more lily your data or mass list is matched to a known peptide. The mowse score cant be tack at face value as its magnitude wont be the only indication that the match is correct. There needs to be a significant difference from the next highest mowse score match to definitively say that the first match is correct. The databases will show the percentage coverage of the known peptide from you mass list as well. The more coverage you have the also more likely to be a correct hit. This type of analysis is called **peptide mass fingerprinting (PMF).**

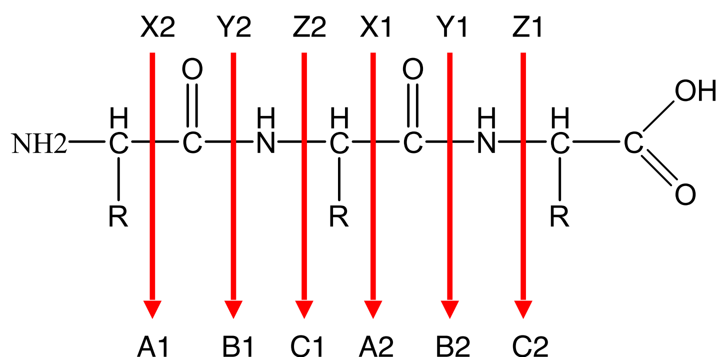
---

24. ABC and XZY ions

ABC and XYZ fragmentation is the way a peptide can be cut up to generate different peptide fragments. Viewd on a spectral graph the original ion will have the highest m/z value and the fragments will be behind it. The fragments will either have the N or C terminus conserved meaning that amino acids are detached from the opposing end.



The fragmentation can occur at different laces but it is normally between amino acids. These fragments are known as B and Y fragments. These fragments will be given a number like B7. The B means the N terminus is conserved and the number relates to how many amino acids are left in the peptide. The fragmentation docent always happen between amino acids though. As you can see from the figure on the right the B locations are between the amino acids but the A and C fragmentation occurs within the amino acids. The same occurs for the XYZ fragmentation But this refers to the C terminus being retained. The most common species are the B and Y ions. The Y ions are more common than the B ions.



Lecture 7

Chapter 7

This chapter will cover:

- MUlti Dimensional Protein Identification Technology MUDPIT.
- Fragmentation issues.

---

25. MUDPIT

The idea behind MUDPIT is to be able to take an organism and pull out all the proteins and digest them. This docent require the need for protein separation and means all the proteins are analysed together.

The first stage of the analyse step is to separate proteins with two trap colours on different valves. It first runs the sample proteins over a trap column based on an ionic property and the second trap column on a hydrophobicity column. You thus split your protein sample into a 90:10 ratio between the two traps. You take a salt and inject it over the cationic exchange column which will loosen the weekly bound ions (the ones based on electrostatic interactions). You then slowly introduce a stronger and stronger salt to release slowly the more tightly bound proteins so that they can flow down to the reverse phase hydrophobicity column trap. You are fractionating bunches of proteins and analysing them one at a time. You usually do around 24 injections which leads to the HPLC analysis which takes about an hour. The advantage of this two port trap system is that you are more accurately able to analyse a mass sample of proteins.

Be aware that some databases will always identify a protein The database returns a list of peptides with a "score" which represents how reliable the data is. Using electrospray – the more proteins the less time and information you gain on each individual protein. So the more proteins you analyse the less reliable your data

Parameter	MALDI-TOF	LC/MS	Comments
Sensitivity	●	●	LCMS has a better separation compared to MALDI/TOF purely because you doing this HPLC separation first. With LC/MC you are separating the ions more so that they hit the detector one at a time but with MALDI you can ion separation where multiple ion types are hitting the detector at the same time.
Dynamic Range	●	●	LC/MS it dosent matter if you have 10,000 of one protein and 2 of another it will separate them but MALDI will get overwhelmed with the higher concentrations. LC/MS can deal with a wider range of concentrations.
Reliability	●	●	LC/MS is more reliable because you getting data about the peptide mass and the MS data.
Coverage	●	●	Here you get more coverage because you are getting less protein suppression and your getting your
PTMs	●	●	Its easier to look for protein transition modification
Time	●	●	MALDI TOF can take minutes to gather dat whereas LC/MS can take hours before getting results.
Cost	●	●	Because LC/MS takes longer you have elevated costs.
● = good	● = OK	● = Bad	

will be because more proteins will be eluted at the same time.

26. Issues with fragmentation

26.1 D and W ions

These ions are like the B and Y ions where a R or functional group is lost from the peptides amino acids. You get D ions if you retain the N terminus of a peptide and you get W ions if you are retaining the C terminus of a peptide. When you get a lost amino acid the amino acid can be charged or ionised.

26.2 Immonium ions

This is when you have retained an N terminus but you have lost a proton from the ammonium creating something like **HN=CH-R.**

26.3 Isobaric amino acids

Isobaric amino acids are amino acids with the same mass. loosing two glycines in one go is the same as loosing one asparagine. The presence of isolucine and leucine is another, To distinguish between these amino acids you would require a D or W ions and these are rare fragmentations. Lysine and glutamine also have a very similar molecular weight of 128.13 and 128.17 respectively. However lysine will be acetylated where glutamine will not to confirm the existence of the amino acid.

26.4 Proline

The aa proline can also stabilise its charge by circulating and bonding its N terms to the free R side chain. This leads to the loss of proline and the amino acid next to it which can be confusing but also used as a way of identifying certain peptides.

Lecture 8

Chapter 8

This chapter will cover:

• Protein modifications

---

27. Protein modification

The most common modifications you can have are:

• Phosphorylation = addition of a phosphate group either on a serene, tyrosine or freonine.
• Glycosylation = carbohydrate residues added on
• Ubiquitonation = Addition of ubiquitine residues
• Farnesylation = addition of fatty acids.

27.1 Phophorylation
Phosphorilation is the most common and has a large number of functions. It is usually carried out to modify the activity of proteins. The enzymes kinases adds phosphates and phosphotases removes them. These eneymzes control the movement of phosphate groups in many secondary signalling systems and thus controls the use of many metabolic activities.
There are two ways to study these phosphate ions. First is to pull out and purify the phosphorilated peptide either through protein specific extraction or recognition of the proteins in an MS machine.

The original study of the phosphate groups was to radio label them with either 32P or 32P-labelled ATP because ATP is the sours of at the phosphate group. You cant study these atoms in a mass spectrometer because you would then have to decontaminate the machine. There are ways around MS analysis though
Immobilised metal affinity chromatography (IMAC) relies on the complexation a metal species to the oxygens on the phosphate groups. Metal in the column to with ions can be held will bind and complex to the sample ions and then elute them off with a high pH. You can therefor alternate the process by incorporating this into a trap column.

Immunoaffinity with Phospho STY can be used in the same way as a 2D gel like the radiolabeled phosphates but this time it is bound to an antibody which will then signal visually with an attached peroxidase like in sandwich ELISA.

Alternativly you could analyse with a mass analyser. There are many ways to analyse the phosphate groups. The first thing to do is derivitise the phosphate groups by converting the sulphur groups by braking the bonds and turning them into S-R groups so they are blocked. And then turn all the phosphate groups into new sulphur groups. Each new sulphur groups now represents a phosphate group. You then react the new sulphur groups with a new chemical group. When analysed now with a CNS analysis you can look for the specific loss of the new chemical to confirm the existence of a phosphate group. You can essentially do the same with phosphopeptides by the loss of the phosphate group (loss of 98Da). The loss of the phosphate will create neutral species and now we know that there is a specific loss we can select for it and pull it out of the sample. You will have to tell the detector to look for more than one loss value though due to the different m/z values possible. Because the peptides are large the charge can be multiplied creating smaller m/z values. so the detector would have to look for a single charge loss and double charge loss and a triple charge loss. If you wanted to specifically identify phosphotyrosine resides though you could do a precursor ion loss analysis (PID) this is because the loss of a two unstable ammonium ions.

27.2 glycosylation
Glycoslyation is used by the addition of a glucose to the peptide to make a glycoprotein. They can act as biomarkeres for cancer or tumour proteins. These glycoproteins represent a majority of surface cell markers over the body so there quite common. there is also an estimation of about 50-60% of secreted proteins to be glycoproteins. There a bit more problematic to create than phosphorilated amino acids. Phosphorilation can occur on three amino acids (serene, tyrosine or freonine) where there is an addition of one phosphate. Glycosylation is more complex because there are many amino acids that can be glycosylated and they can

add to different types of proteins. there can be linear or branched proteins which makes them more difficult to study.

They can be studied either by enrichment methods or through spectral analysis. The first step is to try and purify them by binding carbohydrates on C18 columns. You can get different types of glycan columns by using different types of lectins to try and specify the types of glycans.
You can bind the glycans to specific antibodies that will attach to biotin and cart visual identification like in ELISA. You can use Concanavaline A to bind to the N linked glycans.
You can also look for a specific glycol group in a precursor ion loss and a constant neutral loss.

27.3 Ubiquitination

ubiquitine itself is a protein, when bound to other proteins you get a chain of them. The main function of this is to brake down other proteins. This isn't a great modification to observe because there abundance is low. This is because once ubiquitinated its quickly degraded. You can abstract some of these though. The analysis would require digestion with trypsin and the amnio acids of ubiquitin two things happen. You can look for a speciic change in your spectral analysis. Ubiqutine is usually added to a lysine (oe that trypsin would cleave) within the target amino sequence, so you would genrally get a missed cleavage because it has a ubiquitine added to it. You also get digestion of the ubiquitin peptide which means you get a sequence stuck to your target sequence. You will have a specific digestion of ubiquity which leaves two glycines to the lysine. When you get the fragmentation of the parent peptide, you will get a y ion that has this lys-Gly-Gly mass which is a characteristic loss of a ubiquitinated peptide.

27.4 Farnesylation

A key example of this is the RAS protein with a C18. Usually you need a fatty acid to get a protein to insert into the membrane. You get ferrisylation on the C terminus with R group which then migrates to the membrane. You estimate what peptide you would get with and without the ferrityl group. Trypsin digestion using a single reaction monitoring experiment would be used to find if you get a specific peptide from a specific peptide. You this is basically used to study the analysis of cancers like the mutation of the RAS protein.

Lecture 9

Chapter 9

This chapter will cover:

- COFRADIC
- 2D SDS PAGE gel
- Novel HPLC Isotope-Coded Affinity Tags (ICAT)

---

28. COFRADIC

COmbined FRActional DIagonal Chromatography (COFRADIC) or Diagnoal HPLC. Being able to analyse all the proteins in an organism in one go is hard to accomplish. The main barrier comes from the graduation in concentration amounts. There could be 100mm of one protein and 2nm of another. It would be very hard to observe the smaller concentrations. COFRADIC looks to specifically modify certain proteins within the proteome of an organism to specifically identify them. This technique works in three steps the first is to run a reverse phase HPLC separation on a collection of peptides. The next step is to chemically or enzymatically modify a particular peptide. You then repeat and identical reverse phase HPLC and observe the elimination or reductions of specific peptides. There are four major modifications you can do to the peptides to observe the change in proteomic data:

- methionyl peptides = modification of methionine to create methionine-sulfoxide. Since methionine-sulfoxide is more hydrophilic than methionine, the affected peptides undergo a **hydrophilic shift** and since this shift is quite predictable they can be specifically collected during the secondary separation and are finally used to identify their precursor proteins using automated liquid chromatography tandem mass spectrometry

- cysteinyl peptides = Conversion of cystine into TNB-cystine to make a amino acids more hydrophobic than cystein. These can be separated through HPLC. The TNB group is then removed making the cysteine more hydrophilic and able to collect like with methionyl modification.

- amino terminal peptides (degradomics) = we have changed the sorting chemistry such that now the N-terminal peptides are isolated. Since each protein has one amino terminal peptide, at least in theory, each protein is represented by only one peptide. This can help in separating internal and integral peptides.

- phosphorylated peptides = Briefly, following proteome digestion, phosphorylated peptides are enriched by Fe^{3+}-IMAC and fractionated a first time by RP-HPLC. Phosphorylated peptides are then specifically targeted by a dephosphorylation reaction using a cocktail of ambiguous phosphatases.

---

29. 2D SDS PAGE

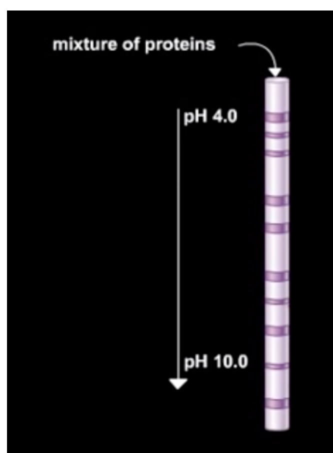
2D-PAGE is a form of gel electrophoresis in which separation and identification of proteins in a sample are done by displacement in 2 dimensions oriented at right angles to one another(orthogonal). This technique is also used to compare two or more samples to find differences in their protein expressions. In this technique proteins are separated by two different physicochemical properties. In the first dimension proteins or polypeptides are separated on the basis of their net charges by isoelectric focusing and in the second dimension they are separated on the basis of their molecular masses by electrophoresis. Because it is unlikely that two molecules will be similar in both properties, molecules are more effectively separated in 2-D electrophoresis than in 1-D electrophoresis.

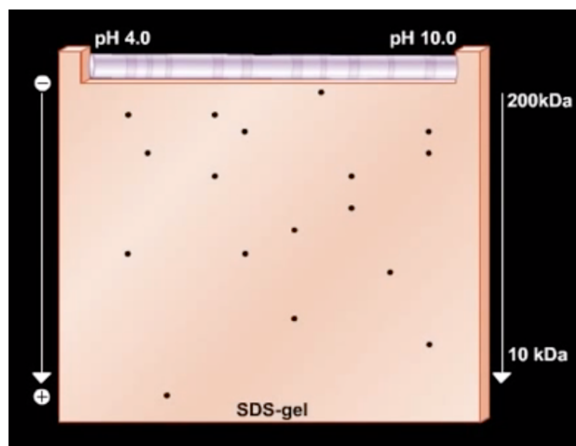
In IEF, proteins are separated by electrophoresis in a pH gradient based on their isoelectric point(pl). A pH gradient is generated in the gel and an electric potential is applied across the gel. At all pHs other than their isoelectric point, proteins will be charged. If they are positively charged, they will move towards the more negative end of the gel and if they are negatively charged they will move towards the more positive end of the gel. At its isoelectric point, since the protein molecule carry no net charge it accumulates or focuses into a sharp band.

Immobilized pH gradients are used for IEF because the fixed pH gradients remain stable over extended run times at very high voltages. The pH gradients of IPGs are generated by means of buffering compounds that are covalently bound into polyacrylamide gels.

The proteins in the focussed IPG strips are uncharged because they are at their pI and so they will not move into the SDS- PAGE gel. So the strips are treated with SDS (sodium dodecyl sulfate), an anionic detergent which denatures the protein by breaking the disulfide bonds and gives negative charge to each protein in proportion to its mass. Without SDS, different proteins with similar molecular weights would migrate differently due to differences in folding, as differences in folding patterns would cause some proteins to better fit through the gel matrix than others. SDS linearizes the proteins so that they may be separated strictly by molecular weight. The SDS binds to the protein in a ratio of approximately 1.4 g SDS per 1.0 g protein (although binding ratios can vary from 1.1-2.2 g SDS/g protein), giving an approximately uniform mass:charge ratio for most proteins, so that the distance of migration through the gel can be assumed to be directly related to only the size of the protein. Proteins may be further treated with reducing agent, such as dithiothreitol (DTT) or TRP(Tributyl phosphine to break any reformed disulfide bonds and then alkalated with iodoacetamide to prevent reformation of disulfide bonds. A tracking dye like bromophenol blue may be added to the protein solution to track the progress of the protein solution through the gel during the electrophoretic run. After electrophoresis the gel is stained to visualize the separated proteins. Commonly used stains are Coomassie Brilliant Blue or SYPRO Ruby or silver stain. Silver stain can detect spots containing proteins less than 1 ng and is the most sensitive non – radioactive protein visualization method.

The method is characterised in the two images below.





Advantage
- Get a visual comparison
- Easily presented and protein purified during an

Disadvantages
- May miss low abundance proteins
- Reproducibility of Gels is questionable
- Practically very time consumingalysis

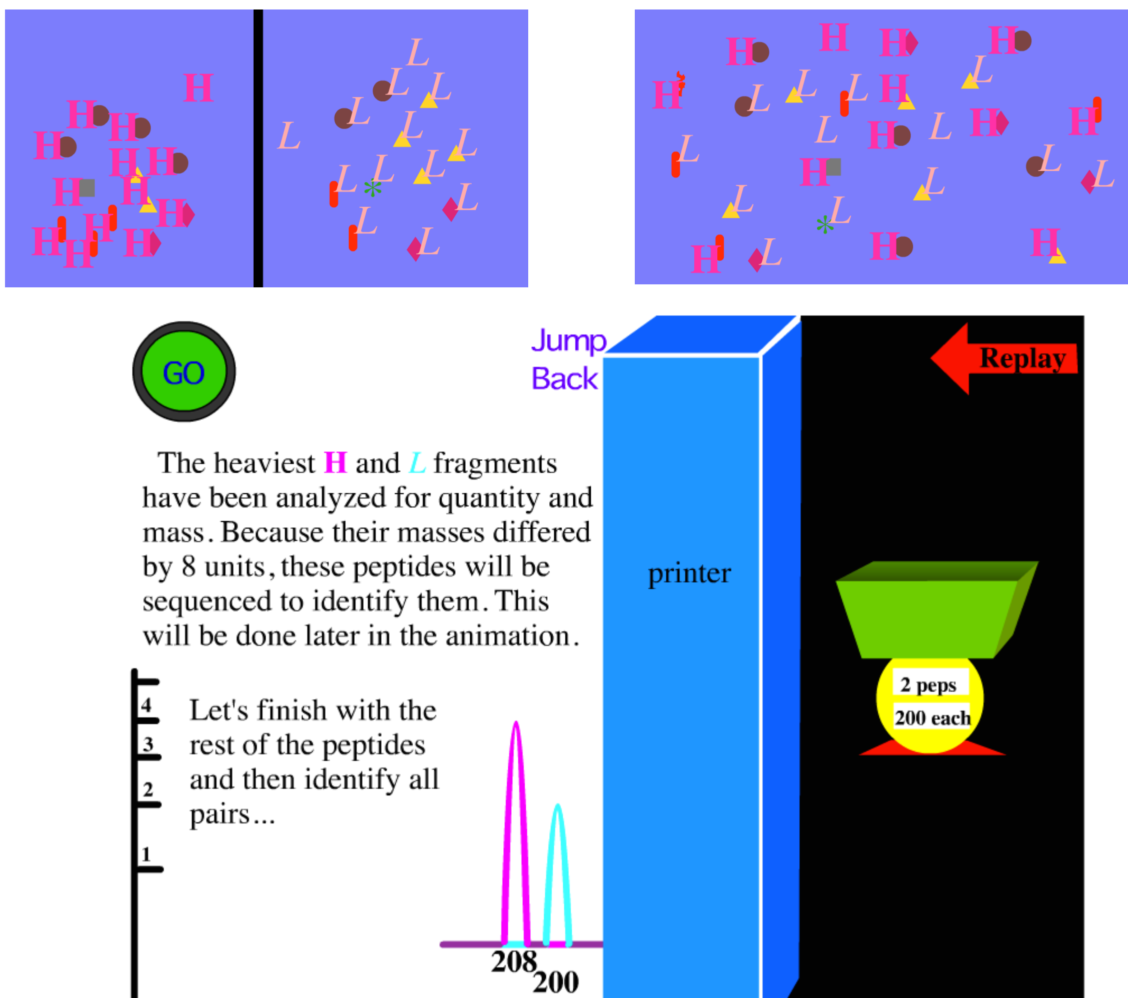
---

29. ICAT

Icat was invented in 1999 and is designed to be a high throughput method for evaluating the protein content of two cell populations. Every protein that contains a cysteine will be measured. I will do ana analysis of yeast to represent the general protocol. This is however a MS way of doing a microarray analysis.

Yeast can be grown in both aerobic and anaerobic conditions. The type of environment the yeast grows in will influence the protein expression. Extraction of these proteins is carried out after the cells have been cultured and digested. The proteome from each cell type is then placed in an ICAT reagent. proteins

congaing cystein will thus have a sulphur group on them which will be attached to a sulphur binding group. Attached to the sulphur binding group will be a tag. This tag will have 8 hydrogen atoms in its structural configuration. One cell type will be placed in a ICAT reagent contain tags with hydrogen, the other cell type will be placed in an ICAT reagent congaing deuterium instead of hydrogen. This will mean that one of the cell types will contain tags that are 8 mass units higher. The two cell types are then mixed together and run through an HPLC analysis. The machine will do an analysis to recognise the mass difference between the two cell type and gather information about the quantities of the different proteins in the two samples. Because the peptides bound to the tags will be digested the peptide sequence can also be analysed through fragmentation MS and the peptides origin can also be determined. This means the whole process indicates the difference in peptide quantities and identifies the proteins the peptides came from.



Advantages

The process is automated so can be done with accuracy and can be highly reproducible.
It gives the relative abundance accurately
and can detach low abundant proteins because of the HPLC separation.

Disadvantages

It produces complicated spectra (twice as many peaks due to different tags
Proteins must contain specific amino acids, There isn't a lot of room for mutations.
The proses it timely and expensive.

Lecture 10

Chapter 10

This chapter will cover:

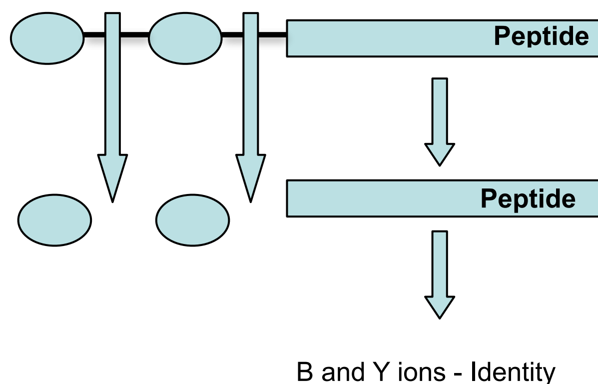
- ITRAQ
- Comparible levels Vs absolute quantisation
- MALDI analysis

---

29. ITRAQ

Isobaric tags for relative and absolute quantitation (**iTRAQ**) are a non-gel-based technique used to quantify proteins from different sources in a single experiment. It uses isotope-coded covalent tags. iTRAQ is used in proteomics to study quantitative changes in the proteome.The ability of ICAT is quite good at determining the difference quantitatively of two different proteins or two samples. The need for more identification of proteins, biomarkers and individual changes requires the need for an analysis of more than one. The limitation of ICAT is that the analysis of light and heavy quantitative approach. This means only two samples can be observed. Multiplexing with ITRAQ has lead to the observation of 8 different biological samples at once.

Labelig chemistry has allowed ITRAQ to develop N-terminal peptide labelling which binds a tag to the terminal amino group or free amino groups in the peptides R groups. The quantitative changes are available with this technique but you are able to observe the experimental differences between the control and treatments cohorts. This is done by the identification of the sample peptide by binding to a specific mass tag. ITRAQ can be done for 8 different samples because you can have eight different balance tags. The tag is made up of a balance group and a reporter group.



The MS machine will accept peptides eluted from a HPLC machine so that the samples are devided up into mass orders. The sample will then enter the MS/MS machine and a variety of specific masses will be selected for by the quadrupole. 1 ion will pass through and be fragmented so that the reporter and balnce group are extracted. The fragment ions will be passed through the second quaripole and if one f the selected reporter tags is identified then the peptides sample location is known. The B and Y ions can then be observed and this allows identification of the peptide through in SILICO analysis.

- Advantages

This process is automated
it gives relative abundance accurately
and can detect low abundances of proteins

- Disadvantages

Its expensive
And it needs suppliers specific MS to carry out an analysis.

29. Comparative levels Vs absolute quantitation

The coverage is defined as the number of proteins found in an experiment that relate to that protein. The peptide coverage is defined as the amount of the protein covered by those peptides. The more you have of a protein the more coverage you will get. You can generally plot the relationships between coverage to estimate protein matching. This is a semi quantitative measure of the levels of the protein. All you do is ask the database how many proteins where associated with each protein. All you need to do is add a internal standard that undergoes the same protocol of your sample of interest. You cant tell how much protein you got from a sample of if you have purified 100% of the sample so the internal standard acts to identify the experimental loss of the sample. If you extracted testosterone you add something to the blood that acts in the same way as the testosterone. This then hoses through the digestion process with it.

The SEA and TSST-1 where investigated because they are the main contributors to food poisoning. they work out what proteins would be produced as part of a protein digestion from the extraction form there organisms (staphylococcus aureus). Then effectively create a nucleotide sequence that would translate to give you that sequence. The created a gene fusion construct. Then give this to your bacteria as a plasmid with radio labels on it. These then act like the protein of interest that acts as an internal standard. This shows you the proteins your getting is correct by digesting it along with the internal standard purification. Taking these from a blood sample and purifying SEA and TSST will thus be proportional to your standard protein. So you can measure the loss during all the processes like temperature changes or spinning down ETC.

30. MALDI analysis of intact proteins.

You stick your protein on a plate and fire a laser at it and get protein mass spectra. Usually you take blood samples and remove things like albumin to clear high concentrations but you can also do MALDI on tissue samples. You can either do direct MALDI to detect which proteins are in your sample or you can do MALDI imaging to see where in the sample the proteins are. Direct MALDI is the protein profiling you do by gathering the time of flight data of the protein fragments and use in SILICO testing to gather the proteins present. The MADI imaging is where you get a tissue and smear a MALDI matrix over it. This then has a mass number of laser sites to impact to gather information about a variety of the tissue sample. This then gives an image of specific protein concentrations over the area of tissue. This allows protein mass signatures in different areas of the tissue like brain tissue to identify the boundaries or uses of the protein in the structures.

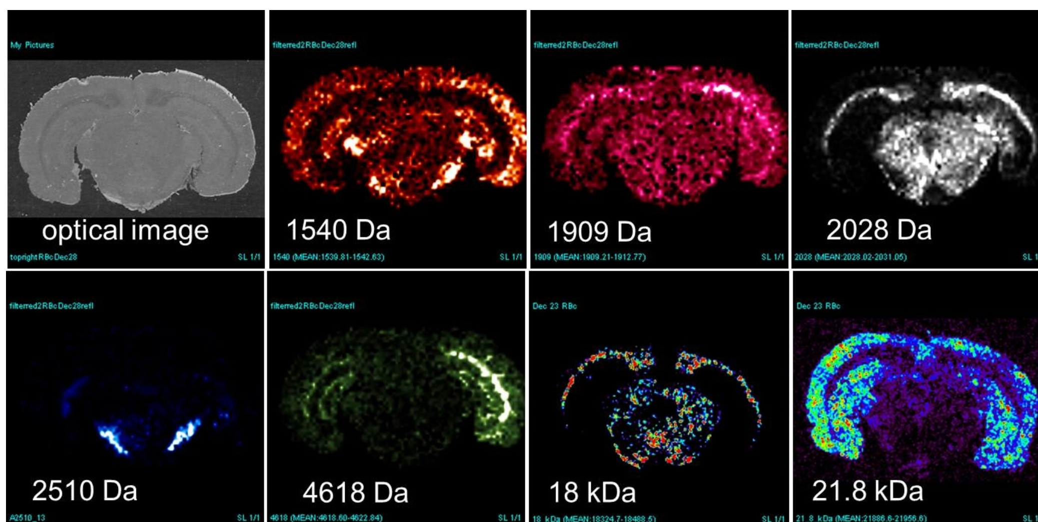
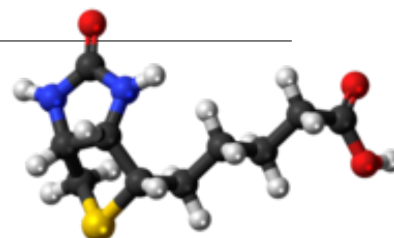


Image of a mouse brain in the optical image and then the concentration changes of the different protein masses through out the section. This is MALDI imagine showing the locations of particular proteins.

The currently being developed to identify bacteria by using the protein profile to analyse the protein content of the bacterial tissue. Theoretically the different bacteria have different genomes therefor have

different proteins and can thus be identified based on there protein composition. This process can work quite well for different species identification but can be quite poor at sub species identification. The change in sub species due to subtle genetic differences can often not be identified by this MS identification technique. Things like antibiotic resistance could be an example of a genetic change that may be hard to identify with this MS technique.
This protein identification technique works well when you have something characteristic like a nucleic protein (proteins that bind to RNA). You can identify proteins that are anchored proteins like the phosphotyrosines.

31. Biotin



Biotin is widely used throughout the biotechnology industry to conjugate proteins for biochemical assays. Biotin's small size means the biological activity of the protein will most likely be unaffected. This process is called biotinylation. Because both streptavidin and avidin bind biotin with high affinity (K_d of 10^{-14} mol/l to 10^{-15} mol/l) and specificity, biotinylated proteins of interest can be isolated from a sample by exploiting this highly stable interaction. The sample is incubated with streptavidin/avidin beads, allowing capture of the biotinylated protein of interest. Any other proteins binding to the biotinylated molecule will also stay with the bead and all other unbound proteins can be washed away. However, due to the extremely strong streptavidin-biotin interaction, very harsh conditions are needed to elute the biotinylated protein from the beads (typically 6M guanidine HCl at pH 1.5), which often will denature the protein of interest. To circumvent this problem, beads conjugated to monomeric avidin can be used, which has a decreased biotin-binding affinity of ~10^{-8} mol/l, allowing the biotinylated protein of interest to be eluted with excess free biotin.

ELISAs often make use of biotinylated secondary antibodies against the antigen of interest, followed by a detection step using streptavidin conjugated to a reporter molecule, such as horseradish peroxidase or alkaline phosphatase.

Proteins can be biotinylated chemically or enzymatically. Chemical biotinylation utilises various conjugation chemistries to yield nonspecific biotinylation of amines, carboxylates, sulfhydryls and carbohydrates (e.g., NHS-coupling gives biotinylation of any primary amines in the protein). Enzymatic biotinylation results in biotinylation of a specific lysine within a certain sequence by a bacterial biotin ligase. Most chemical biotinylation reagents consist of a reactive group attached via a linker to the valeric acid side chain of biotin. As the biotin binding pocket in avidin / streptavidin is buried beneath the protein surface, biotinylation reagents possessing a longer linker are desirable, as they enable the biotin molecule to be more accessible to binding avidin/streptavidin/Neutravidin protein. This linker can also mediate the solubility of biotinylation reagents; linkers that incorporate poly(ethylene) glycol (PEG) can make water-insoluble reagents soluble or increase the solubility of biotinylation reagents that are already soluble to some extent.

The most common targets for modifying protein molecules are primary amine groups that are present as lysine side chain epsilon-amines and N-terminal α-amines. Amine-reactive biotinylation reagents can be divided into two groups based on water solubility. N-hydroxysuccinimide (NHS) esters have poor solubility in aqueous solutions. For reactions in aqueous solution, they must first be dissolved in an organic solvent, then diluted into the aqueous reaction mixture. The most commonly used organic solvents for this purpose are dimethyl sulfoxide (DMSO) and dimethyl formamide (DMF), which are compatible with most proteins at low concentrations. Because of the hydrophobicity of NHS-esters, NHS biotinylation reagents can also diffuse through the cell membrane, meaning that they will biotinylate both internal and external components of a cell.

Purification
The biotin tag can be used in affinity chromatography together with a column that has avidin (also streptavidin or Neutravidin) bound to it, which is the natural ligand for biotin. However, harsh conditions (e.g., 6M GuHCl at pH 1.5) are needed to break the avidin/streptavidin - biotin interaction, which will most likely denature the protein carrying the biotin tag. If isolation of the tagged protein is needed, it is better to tag the protein with iminobiotin. This biotin analogue gives strong binding to avidin/streptavidin at alkaline pH, but the affinity is reduced upon lowering the pH. Therefore, an iminobiotin-tagged functional protein can be released from an avidin/streptavidin column by decreasing the pH (to around pH 4).

Detection

This tag can also be used in detection of the protein via anti-biotin antibodies or avidin/streptavidin-tagged detection strategies such as enzyme reporters (e.g., horseradish peroxidase, alkaline phosphatase) or fluorescent probes. This can be useful in localization by fluorescent or electron microscopy,[9] ELISA assays, ELISPOT assays, western blots and other immunoanalytical methods. Detection with monovalent streptavidin can avoid clustering or aggregation of the biotinylated target.

Chapter 11

This chapter will cover:

- Single nucleotide polymorphisms (SNP's)
- Ion mobility

---

32. SNP's introduction

When you look at proteomics basically all your looking at is the gene expression. You can use proteomics to look at the gene expression levels or the gene products. You could take a single cell and you could look at the expression of proteins and the mRNA (transcriptomics). You can then look at the sum of them at integration. There are mRNA strands that instant translated and there are proteins hows mRNA is degraded. You cant effectively look at one with out the other then which is why both are used. Using ITRAQ you can look at the metastasis in biopsies and serum. They took healthy patients biopsies (non malignant) they took non progressing bipsires (benighn or low malignancy) and then progressive (malignant) biopsies and measured the cell protein expression in these cells and found the difference in the level in proteins that can effectively catalogue the progress of the disease.
You then go back and verify this with another identification system such as chemical sating and looked at the level of the tags in the tumours to determine the level of the tumor.

This analysis can also be used to measure protein expression in plants in different growing conditions. The ratio changes can be related to a score due to the degradation of certain proteins and which were consistant. You can look for single proteins from single sources but you can also look at complex proteomes called metaproteomics. This is usually the protein analysis from the environment which includes many organisms. You extract the proteomes from these organisms and come out with a environmental proteome which is characteristic of that environment. This type of study can be done to look at the difference the change in things like different soil types and contaminated soils to look at the difference or in areas of flooding and how this changes the environment. This can even go into studying the changes of forests in the past. You can analyse the species like bacteria and fungi. The change in these organisms can be characteristic of the change in the forrest.

An organism of interest is gliaficky a particular which grows over monuments. It has been shown that they work together called quorum sensing. You can use the proteome to create a artificial one which can be used in the lab. Exchanging the nature of the proteome can affect the proteins

---

33. Ion mobility

Ion-mobility spectrometry (**IMS**) is an analytical technique used to separate and identify ionized molecules in the gas phase based on their mobility in a carrier buffer gas. Though heavily employed for military or security purposes, such as detecting drugs and explosives, the technique also has many laboratory analytical applications, recently being coupled with mass spectrometry and high-performance liquid chromatography. IMS devices come in a wide range of sizes (often tailored for a specific application) and are capable of operating under a broad range of conditions. Systems operated at higher pressure (i.e. atmospheric conditions, 1 atm or 1013 mbar) are also accompanied by elevated temperature (above 100°C), while lower pressure systems (1-20 mbar) do not require heating.
The ion mobility MS is used in mass spectrometry where a the ions move through the drift cell where they encounter a gas to slow them down. The ideal gas is something that wont react with the compounds. The rate of movement through the gas is to do with the size and shape of the molecule rather than mass. With this approach you can identify the shape change of the proteins. You identify subtle changes i the shape of things without the change of its overall sequence. You can also use it look at a bacteria Grawl. The mass of this bacteria is huge (50Kba) but it becomes more compact when bound to certain proteins. The change in drift time can be used to identify its state.

34. SNP's continued

An SNP is a single base change in the DNA sequence which may or may not result in a protein sequence. There are generally 200million SNPs in the human genome that account for al the variations in humans. They can change looks, personality and disease susceptibility and radiation sensitivity. SNP's can have either one of two effets They can have no net effect on the protein sequence. Alternatively they do and the rise in amino acids is different and this can affect the protein. There are two outcomes of this. the SNP can have some little effect on the protein or they can have no effect still. MALDI can be used to study the SNP's to study the primary extension to detect the difference. You first do PRC to amplify the sequence with a specific extension primer and amplify the DNA in a master mix missing one of the nucleotides. The DNA will extend but wont be able to add one nucleotide and will terminate the extension. The identification of an SNP will be the continuation from the primer by different distances. If you are analysing a SNP then you can creat a primer that will add right next to the SNP site. it can either be extended by a couple of base pairs and stop at the SNP because of the missing nucleotide from the master mix. If there isn't a SNP the it will continue down the sequence until the next time it needs the missing nucleotide which will be 2 or three more down the sequence. Its the difference in these extended sequences that will determine the presence of a SNP or not. You can also tell if this SNP is homozygous or heterozygous by the quantity of the SNp being analysed in the MS analysis. A SNP analyse usually involves sequencing swell which can be expensive and its not always that reliable. The MS analysis is an alternative to this sequencing which is cheap and easy in comparison although you would normally use sequencing to back up your findings. Finally as long as you design your extension primer a MALDI can discriminate between the peaks of the SNP's to tell what SNP's are present in seven different genes in one go. This can be done as the resolution of the AMLDI machines are accurate enough to determine the difference between closely related masses.

Summary
In conclusion the study of proteins via SNP's with changes that affect the amino acid sequence and the structural configuration of the protein can be identified and measured with ion mobility analysis. Proteins with SNP's can also be measured for changes with MALDI as can the DNA sequences when applying SNP analysis with missing dNTP's in the master mixes. Most of the analysis via spectroscopy techniques is accurate but still needs the sequencing of the DNA and the presence of the mRNA to confirm its mutation.

Lecture 12

Chapter 12

This chapter will cover:

• Small molecule quantitation

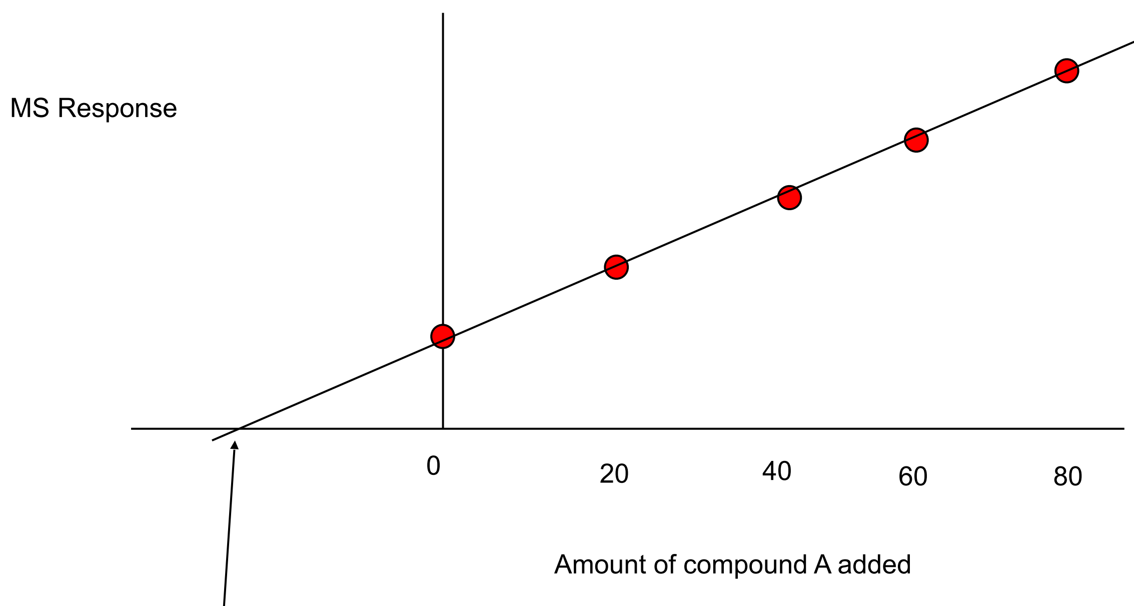
---

35. Small molecule analysis

Small molecule analysis usually looks at the quantisation of a compound that may be present in small amounts. This may include the the drugs that atheleats use to increase performance or things like the presence of drugs on money (cocaine). If your doing quantitation there are three main mechanism to ensure quatitation is accurate.

• Standard Addition experiment
• External standard
• Internal standard

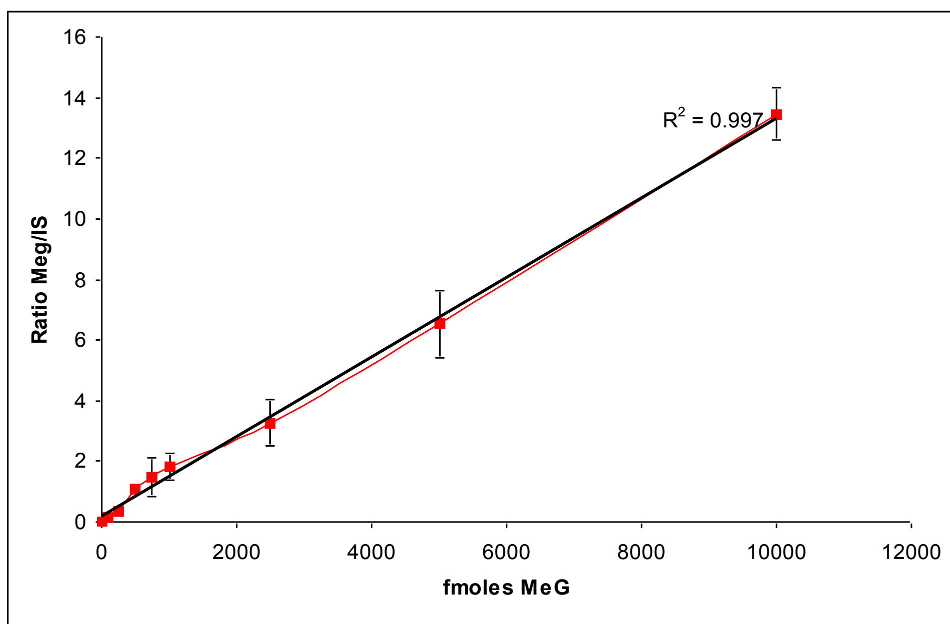
Standard addition experiment is the simplest technique, it introduces in to a mass spectrometer an an HPLC or MALDI plate and you detect the response. You take your sample and split it in to four aliquots of equal volume. To the first you add 100uL of water and you MS analyse it. The next aliquot you add 100uL of water with 20pg of X into your sample and analyse that. The third one has 40pg of X and the fourth has 60pg of X. You get a spike sample from these and an increased peak sample from the four aliquotes. Doing this aids you increasing a standard for your device like an MS devise as ling as its concentration sensitive. The elevation in the material.



You have to make sure theres a linear relationship over your standard samples. If you have a very small sample then you have to do five separate analysis. But in principle it allows for very accurate sample analysis. The key advantage is that any background from the background material docent matter because you have the same concentration in each sample.

Internal or external standards these are things that usually behave in the same way as your sample of interest. An internal stander is something you add to your standard before you purify the sample f interest. An external is one that you would add after. you usually only use one of these standards. The external stander usually is used to account for the reduction in signal due to the contamination of the MS. This will also fall in repeat samples like it may do due to the reduction of the samples detection after the MS machine

gets dirty. The ratio to your compound in the sample will stay the same though. This thus acts as an indicator for. The internal standard is a compound that is added to your sample before the purification so that you can monitor the effects on you sample of interest in every technique. This also counts for the loss in the purification and analysis whereas the external standard only measures the analysis.



Normaly a small an internal standard will be isotopically labeled. This would be done with two N15s in your. This has the benefit of not altering the peak of your sample of interest. Usually an internal standard is a radiolabeled copy of the compound of interest. This will then extract in the same way. There losses will be identical but when put in the MS the mass of your compound of interest will be different because of the radioisotope. This means the chemical reactivities will react in the same way. This is just one of the perimeters of the FDA.

Selectivity = You must take between 6 and 10 blank plasma samples to make sure that there are no peaks in there that will affect your assay.

Sensitivity = (Limit of detection / limit of quatitation) You study the signal to noise ratio is at signal to one. Your sample must be 3x your background. The quantitation two or three times the LOD this will be measured over days and week.

Liniarity of standard curve = its a requirement that the standard curve of your standard is linear.

Recovery = Making sure you are purifying the sample in the correct way

Accuracy = you then have to measure 4 Quality control experiments to make sure that your sample being detected is actually what you wanted.

Sample stability = Usually you want to make sure the sample docent degrade over time to make sure your analysis is accurate.

36.Metabolomics

Transcriptomics looks at the peel of mRNA, proteomics looks at the level of the proteins and metabolmics looks at the levels of metabolites. These should all be linked the metabolites are products of the enzymes functions of the cell and the enzymes are proteins. Metabolomics seeks to get a view a profile of the cell in an unbiased manner without biased to a specific group of metabolites. It looks for changes in metabolites based on a stimulus or environmental cues wether it be a change in fatty acids. genomics is the nalysis of the DNA sequence,transcriptomics is the analysis

of what is transcribed, proteomics is the analysis of whats translated although it wont relate to post translational modifications and finally metabolomics is the analyse of how all of this affects the cell. You can subdivide metabolomics into

- Exometabolomics = Things that are secreted out of the cell. You can cutler cells in a medium and analye the medium there growing in to devise what there growing in.

- Endoetabolimics = Things that remain on or in the cell.

Fingerprinting or foot printing is looking at metabolics as ana analysis technique foot printing looks at the exomome and fingerprinting looks at the endome. When looking at these cells you want to look at what there doing at that exact point in time which requires quenching which is effectively stopping the metabolism. Targeting metabolism is looking at specific things like a mouse heart then you would know effectively whats there.

---

35. Targeted metabolomics

This can include looking at specific components of the metabolome based on the preconceived idea that a stimulus, environmental or genetic change may affect the specific metabolites such as sugers or fatty acids in the membrane. One direct area of targeted metabolomics is lipidomics which is essentially the analysis of phospholipids. This utilises the faculties of th MALDI machines being able to use lasers to sublimate tissue samples to directly analyse the surface proteins of the cells. These phospholipids can be doth accurately identified based on the accurate mass analysis of the MALDI TOF instruments. You can also use CNL to determine targeted samples like particular phospholidids to determine the head group that will be lost when fragmented.

Lecture 12

Websites:

http://www.astbury.leeds.ac.uk/facil/MStut/mstutorial.htm

Review Papers:

http://www.sciencedirect.com/science/article/pii/S0031942204001645

http://www.sciencedirect.com/science/article/pii/S0031942211000586

http://jcb.rupress.org/content/190/4/491.long

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2642903/

Table of amino acid residues

Symbol	Structure	Mass (Da)
Ala A	-NH.CH.(CH_3).CO-	71.0
Arg R	-NH.CH.[(CH_2)$_3$.NH.C(NH).NH_2].CO-	156.0
Asn N	-NH.CH.(CH_2CONH_2).CO-	114.0
Asp D	-NH.CH.(CH_2COOH).CO-	115.0
Cys C	-NH.CH.(CH_2SH).CO-	103
Gln Q	-NH.CH.($CH_2CH_2CONH_2$).CO-	128.1
Glu E	-NH.CH.(CH_2CH_2COOH).CO-	129.0
Gly G	-NH.CH_2.CO-	57.0
His H	-NH.CH.($CH_2C_3H_3N_2$).CO-	137.1
Ile I	-NH.CH.[CH.(CH_3)CH_2.CH_3].CO-	113.1
Leu	-NH.CH.[$CH_2CH(CH_3)_2$].CO-	113.1
Lys K	-NH.CH.[$(CH_2)_4NH_2$].CO-	128.1
Met M	-NH.CH.[$(CH_2)_2$.SCH_3].CO-	131.0
Phe F	-NH.CH.(CH_2Ph).CO-	147.1
Pro P	-NH.$(CH_2)_3$.CH.CO-	97.1
Ser S	-NH.CH.(CH_2OH).CO-	87.0
Thr T	-NH.CH.[CH(OH)CH_3).CO-	101.0
Trp W	-NH.CH.[CH_2.C_8H_6N].CO-	186.1
Tyr Y	-NH.CH.[(CH_2).C_6H_4.OH].CO-	163.1
Val V	-NH.CH.[$CH(CH_3)_2$].CO-	99.1

Lecture 12