Protein Mass Spectrometry for Systems Biology

Introduction

Much like a biological system, protein mass-spectrometry (MS) relies on the well-coordinated, intricate interaction of many components (sample prep, isotopic labeling technologies, ionization methods, mass filters, detectors, databases, analysis software etc.) (Figure 1). For each of these components multiple alternatives exist and while these components are to some degree modular, not all combinations of them work together. For example, certain ionization methods will only work with certain mass-analyzers. Or certain techniques of quantification require the use of certain sample preparation methods etc.

Therefore, the description given in this hand out does not even attempt to provide an overview of mass spectrometry. Instead, the goal is to give an idea of the basic concepts and to give a few concrete examples of "typical" MS experiments and instruments that illustrate these concepts.

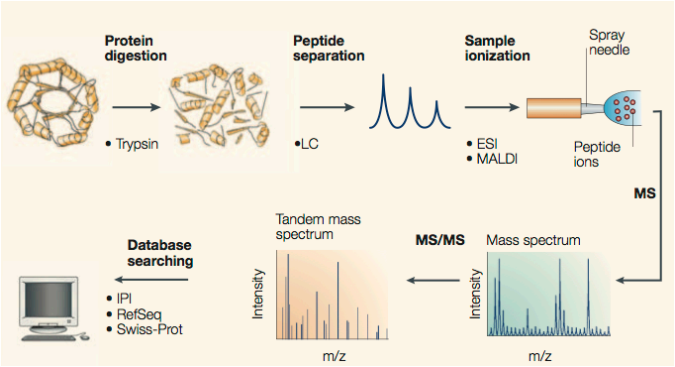


Figure 1 *Schematic representation of a modern protein mass spectrometry experiment. The protein digestion and peptide separation experiments are part of the sample preparation stage. Sample ionization, the two mass separation (MS) stages, the intervening fragmentation step (not shown) and the detection take place inside the mass spectrometer. The data is then analyzed with the help of elaborate computer algorithms and reference databases.*

Fundamental operating principle and components of a mass spectrometry

The fundamental goal of mass spectrometry is determining the mass of individual molecules by measuring how electrically charged versions of these molecules (ions) move when exposed to electric or magnetic fields in a vacuum.

A mass spectrometer (figure 2) consists minimally of an **ion source**, which deposits an electrical charge on the molecule and transfers it from solution into the gas phase, a **mass analyzer** which separates different ions based on their mass to charge ration (*m/z)* and a **detector** counting the ions passing the mass analyzer (Figure 2).

The mass filters/analyzers are constructed in such a way that ions of different *m/z* are separated and hit the detector at different time points or under different electromagnetic conditions of the mass analyzer (scanning). By recording the signal on the detector while scanning across a range of mass-to-charge (m/z) ratios a mass spectrum, that is to say the relative abundance of ions with different mass-to-charge (m/z) ratios, can be recorded.

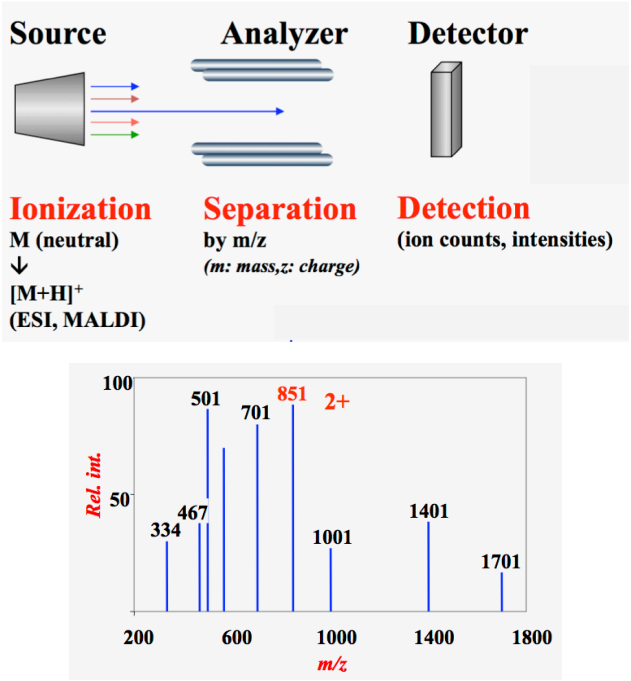


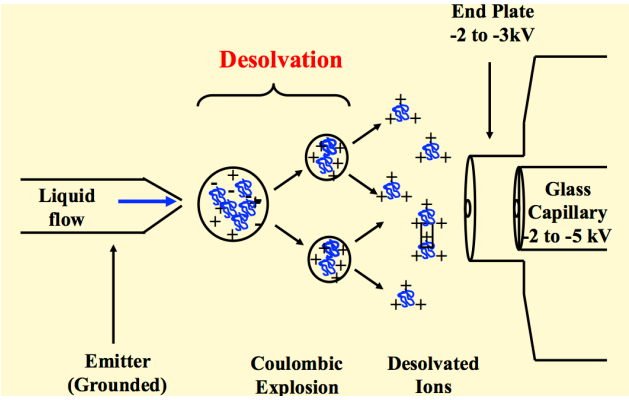
Figure 2 *Simple scheme of a mass spectrometer (**a**). The source deposits a charge on the sample molecules, transfers the resulting ions into the gas phase and (unless a separate accelerator exists) accelerates them in a continuous molecular beam. This beam passes through the mass filter/analyzer section, which separates ions entering the system based on their mass-to-charge (m/z) ratio. There are several physical implementations of mass analyzers. The mass filter/analyzer is adjustable and can scan the window of m/z ratios that are permitted to pass by across a wide range of values. By recording the intensity of the signal at the detector as a function of the m/z ratio one obtains a mass spectrum (**b**).*

Ion sources

Biological molecules usually exist in aqueous solution. In order to analyze them in a mass spectrometer, these molecules need to be charged, de-solvated and transferred into the gas phase.

These three steps are carried out within the ion source. In modern mass spectrometers two main types of ion sources are used: Electrospray ionization (ESI) and matrix assisted laser desorption & ionization (MALDI). Many of the mass spectrometry applications in systems biology use ESI ion sources (Figure 3) and their physical principle is discussed here.

ESI ion sources spray a very fine stream of liquid out of a nozzle that is at a strong positive electrostatic potential relative to a plate at the opposite side of the injection chamber. As a result, the droplets leaving this nozzle carry an excess positive charge. On the way to the mass analyzer (which is in high vacuum) the droplets pass through a zone of dry gas that is frequently heated, letting the solvent in these droplets evaporate very quickly. Thus, the charges on the surface of these droplets are concentrated to the point where a "coulombic explosion" shatters the droplets into smaller droplets, which again shrink due to evaporation and explode again until the solvent has evaporated completely and only the charged ions of the sample molecules remain. Because the solvent molecules are neutral, they are not accelerated by the electrostatic field and can simply be removed by the vacuum pumps.



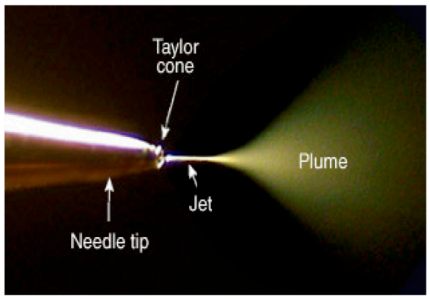


Figure 3 *Schematic representation (**a**) of the processes in an electro spray ion source (ESI). A stream of positively charged droplets is sprayed from a fine nozzle into a sample chamber. The droplets are accelerated towards a plate at the opposite side of the chamber, which is at a strong negative potential. De-solvation of the droplet due to evaporation leads to repeated coulombic explosions of the droplets until only the charged sample molecules remain, which are pulled into the flight path of the mass spectrometer. (**b**) Shows a photograph of the inside of an actual electrospray ionsource. The electrostatic repulsion between the droplets leads to a splaying of the ejected droplets into a cone of droplets.*

Mass filters/analyzers

A mass filter is a device that separates accelerated ions. It only lets them pass if they have a particular mass-to-charge (m/z) ratio that has been specified by the experimenter. Several different types of mass filters exist that are based on rather different physical principles and that have different performance characteristics (more information below). As an example of the physical principles that can be employed in a mass filter the operating mode of a quadrupole mass filter is discussed in Figure 4.

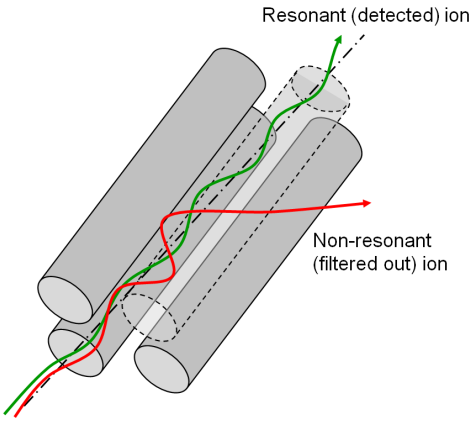


Figure 4 *In a quadrupole mass filter, accelerated molecular ions pass through a channel surrounded by four metallic rods, to which a quickly oscillating electric field is applied. The resulting oscillating electrostatic field sets the ions in the channel on an oscillating path. Unless the oscillation of the field and the induced oscillation of the ion resonate perfectly, the ions will veer off path, collide with the chamber walls and thus be filtered out. The ion's frequency of oscillation is determined by its mass to charge ratio. So by controlling the frequency of the quadrupole's oscillations it is possible to select the mass to charge ratio of the ions that are able to pass this filter and all other ions are rejected. If the field is turned off, the filter is deactivated and all ions are able to pass. If a different setting of the quadrupole is applied, ions with a different m/z are selected. If the setting of the quadrupole conditions changes over time, at every setting and thus at every time point ions of a different m/z are selected (scanning).*

Detectors

The final element in a mass spectrometer is the detector which records the arriving ions. This is done either by the ions physically striking the detector surface or by the ions passing by the detector surface and the detector sensing the field of the passing ion. As the actual number of ions arriving at the detector is rather small, the signal needs to be amplified considerably (e.g. by an electron multiplier) before it can be measured.

MS / MS spectrometers

Many modern mass spectrometers possess two or more mass filters that are arranged in series along the flight path of the ions. Between these two mass filters the ions pass through a **fragmentation stage**. During this phase, molecular ions are broken down into smaller pieces.

One of the most popular fragmentation stages is a collision induced dissociation stage. In such a CID stage the fast-moving molecular ions are allowed to collide with trace amounts of neutral gasses (often helium, argon or nitrogen). Part of the kinetic energy of the collision is converted into intra-molecular vibrations so strong they can disrupt chemical bonds within the molecular ions. The resulting fragments, provided they carry a charge, continue on their flight path while the neutral fragments are removed by the vacuum pumps together with the collision gasses.

In a typical MS/MS experiment (Figure 5) the first mass filter selects ions of a particular m/z ratio. The selected ions enter the fragmentation stage where they break down into fragments that are analyzed in the second mass filter. By coordinating the scanning patterns of the first and the second mass filter it is possible to select each of the sample ions, to fragment them and to record a mass spectrum of the resulting fragment ions. The fragmentation pattern of a sample ion is often very characteristic and, together with the sample ion's m/z value, can be used to identify the chemical identity of the sample ion.

In practice the number of sample ion types that arrive at the first mass filter is so large that it is not possible to record fragmentation spectra for each one of the sample ions. The optimal strategy for selecting the sample ions that are fragmented and analyzed further depends very much on the overall scientific question to be answered. This will be discussed in greater detail during the lecture.

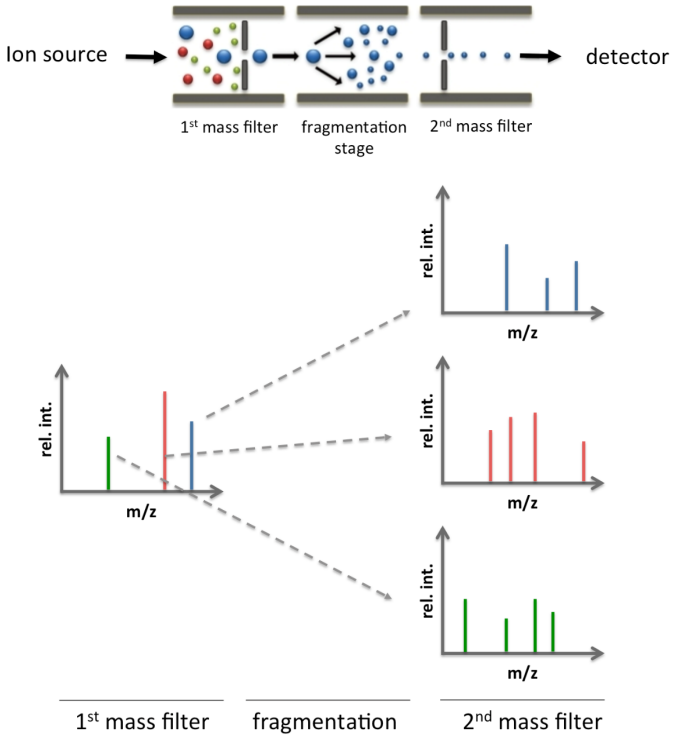


Figure 5 *Operating principle (**a**) of an MS/MS mass spectrometer consisting of two mass filters with an intervening fragmentation stage. The first mass filter can select, based on their m/z ratio, which sample ions enter into the fragmentation stage. The second mass filter is then used to filter the generated fragment ions according to their m/z ratio. (**b**) An MS/MS instrument therefore allows the selection of a sample ion according to its m/z ratio in the first mass filter and a further analysis of this ion by breaking it apart in the fragmentation stage and analyzing the resulting fragment ions in the second mass filter.*

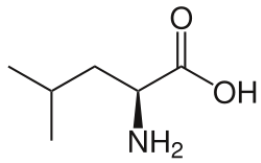
Features of Mass Spectra

As discussed above, a mass spectrum is a plot of the relative intensity of the detector signal as a function of the mass-to-charge ratio (m/z) of the ions generating that signal. It is fundamental to the understanding of mass spectroscopy to realize that one peak in the spectrum does **not** correspond to one molecular entity (e.g. an amino acid or a peptide). Multiple chemical entities may contribute to one peak and one molecular entity will typically generate multiple peaks in a mass spectrum.

The mass spectrometer measures the mass of a molecular species and this mass is determined by the sum of the masses of its atoms. Therefore, every molecular species with the same atomic composition (and charge) will contribute to the same peak in the mass spectrum, regardless of how the atoms in the molecule are arranged. Molecular species that have exactly the same mass are called isobaric species. The classic

example of isobaric amino acids is leucine and isoleucine (Figure 6), both of which contain 6 carbon, 13 hydrogen, 1 nitrogen and 2 oxygen atoms, but have distinct chemical structures. Equivalently, the peptides Asp-Tyr-Pro-Lys-Glu-Thr and Tyr-Pro-Lys-Glu-Thr-Aps are isobaric because they have the same amino acid composition.

a



b

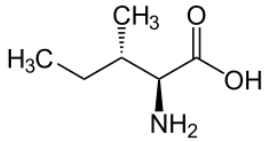


Figure 6 *Example of two isobaric amino acids leucine (**a**) and isoleucine (**b**). Both molecules have the same atomic composition ($C_6H_{13}NO_2$) and therefore have the same mass. Yet they are chemically distinct and have different molecular properties.*

Multiply charged ions

A single type of molecular entity in the sample can generate multiple peaks in the mass spectrometer by adopting multiple charge states. In the ionization phase one of two identical sample molecules may receive a single charge while the other one may receive two charges (doubly charged ion) or even three charges (triple charged ion). Because the mass spectrometer actually measures the mass-to-charge ratio (m/z) and not the mass, the doubly charged ion will appear in the mass spectrum at roughly half the m/z ratio of the singly charged ion, even though they originate from identical sample molecules. The actual m/z recorded for a doubly charged ion is $[M+2H]^{++}/2$.

Isotopes

Isotopes are atoms that contain the same number of protons and electrons, but differ in their number of neutrons. They therefore display the same chemical properties, but have different masses.

A list of naturally occurring isotopes for elements commonly found in biomolecules is shown in Figure 7. The isotopes of carbon are particularly important because carbon is present in virtually all biomolecules and it possesses a relatively common isotope. Approximately one, in a hundred carbon atoms has an atomic mass of

Atom	Mass	rel. Abund (%)
carbon	12.000	98.900
	13.003	1.100
hydrogen	1.008	99.985
	2.001	0.015
nitrogen	14.003	99.630
	15.000	0.37
oxygen	15.995	99.760
	17.999	0.200
sulfur	31.972	95.020
	33.968	4.210

Figure 7 *Table of the relative abundance and atomic mass of different elemental isotopes.*

13 instead of 12. ^{13}C carbon atoms have the same chemical properties as ^{12}C carbon atoms, they are therefore randomly integrated into biomolecules during natural biosynthesis.

So even if all molecules of a group are chemically identical, the group will contain some molecules in which one, two, three or more ^{12}C carbon atoms have been replaced by a ^{13}C carbon atom, resulting in those molecules having a higher mass. Therefore, even chemically pure compounds will consist of a mix of molecules with different masses. In a mass spectrum this is visible as a so-called isotope ladder (aka isotope distribution) of multiple peaks shifted relative to one another by one mass unit (Figure 8). The multiple ion species appear because in a molecule with a number of carbon atoms the number of ^{13}C carbon atoms follows a Poisson distribution.

Besides leading to relatively crowded spectra, the presence of these isotope ladders pose stringent

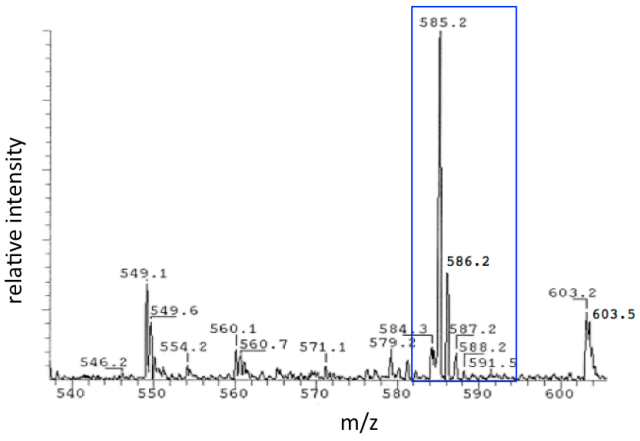


Figure 8 *Zoomed in view of a peptide mass spectrum showing the presence of a typical isotope series of mass peaks (blue box). Note how the masses of the peaks increase by one mass unit and decrease in intensity. These "isotope" peaks are caused by variants of the molecule that contain one, two, three etc. ^{13}C atoms.*

requirements on the minimal resolution of a mass spectrometer. If the resolution of a mass spectrometer is not sufficient to resolve isotope ladders the peaks fuse and the resulting aggregate peak will have a mass value that is hard to interpret (Figure 9).

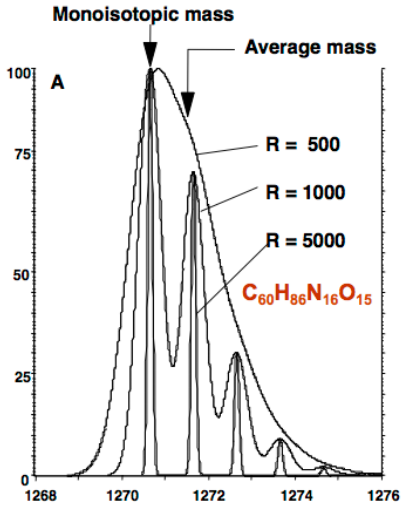


Figure 9 *An isotope ladder in the mass spectrum of a biomolecule with the sum formula of $C_{60}H_{86}N_{16}O_{15}$ shown at three different resolutions (R=5000, 1000 and 500). As you can see, if the resolution is comparable to or greater than the m/z value of the ions being analyzed, the individual peaks of the isotope ladder can be separated. At lower resolution the peaks of the isotope ladder fuse to an aggregate peak. Neither the peak mass nor the average mass of the resulting aggregate peak correspond to the m/z of any of the underlying ions and the measured mass value becomes difficult to interpret.*

Resolution

The resolution of a mass spectrum has such a great impact on its interpretability and usefulness, that it seems reasonable to briefly review the concept of resolution (and the related concept of accuracy) and to compare the performance of different mass filter types commonly used in mass spectrometry.

In mass spectroscopy the resolution of an instrument is customarily defined as the full-width at half maximum (FWHM) of an isolated peak in the spectrum (Figure 10).

So the resolution typically depends on the mass of the ion being measured. The higher the m/z value of the ion, the lower the resolution in absolute m/z terms. It is therefore usually reported as the relative resolution R, which is calculated as the FWHM of the peak (in m/z) units divided by the m/z value of the peak's center.

This leads to another important parameter which is the accuracy of the m/z value. The question is how accurately the center of a peak in the mass spectrum corresponds to

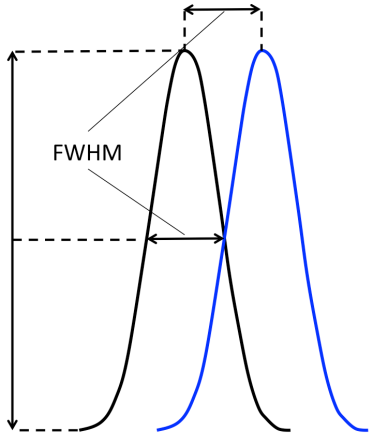


Figure 10 *The resolution of a mass spectrometer is typically defined as the full-width at half the maximal peak height (FWHM). Two peaks that are separated from one another by that distance are considered to be distinguishable.*

the actual m/z value of a measured ion. The accuracy is reported in ppm. An accuracy of 100 ppm, for example, indicates that the m/z value of an ion with m/z = 1000 can be measured with an accuracy of 100 x 1000/1'000'000 = 0.1 m/z units.

Both the resolution and the accuracy of a mass spectrometer are determined by the characteristics of its mass filter. Figure 11 shows a table of the performance of different types of mass filters.

Mass Analyzer	Mass Accuracy	Resolution* (FWHH)	Scan Rate	m/z Range
TOF	Good / very good 1 – 25 ppm	High 10,000 – 40,000	Fast / very fast 10s of ms or less	Very wide > 20000
Ion trap	Moderate 100-1000 ppm	Low 1,000	Fast 10 – 100 ms	0 – 4000 (usually ~2000)
Quadrupole	Moderate 100-1000 ppm	Low 2,000	'Normal' ~1 s	0 – 4000 (usually ~3000)
Orbitrap	Excellent < 1 - 5 ppm	Very high 60,000	'Normal' ~1 s	100 – 4000 (usually ~2000)

Figure 11 *Performance characteristics of different mass filters commonly used in mass spectrometers.*

Sample Preparation

One of the central tasks of mass spectrometry in systems biology is the analysis of the protein composition of living tissues or cell cultures. Particular attention is payed to the question of how this composition changes in response to perturbations.

The proteins need to be extracted from the tissue and prepared for analysis in the mass spectrometer. Standard steps for extracting proteins from complex samples are the disruption of the tissue by physical means (blender,

cryo-pulverization etc.) followed by chemical extraction steps using detergents and salts.

In many cases one is only interested in the overall protein composition of the sample and not in the physical interactions between different proteins, so the proteins can be denatured (i.e. unfolded) very early on in the sample preparation process. In other cases, such as the affinity purification experiments you have learned about previously, it is necessary to keep the tertiary and quaternary structure of the proteins intact until after the affinity purification step.

Either way, protein samples are always denatured prior to analysis in the mass spectrometer. This involves the unfolding of the protein and the chemical reduction of any disulfide bonds within. To prevent these from reforming, the cysteine residues of the protein are alkylated.

Proteins are digested into peptides prior to MS analysis

While it is technically possible to analyze the masses of intact proteins in a mass spectrometer, this is only done rarely and if the research question demands so.

For as the physical and chemical properties of proteins vary so much from one protein to the next, it is difficult to develop one protocol that can be applied consistently across a wide range of different proteins.

Oligopeptides, on the other hand, tend to have relatively similar physicochemical properties that are independent of the peptide's exact sequence. It is therefore possible to develop general protocols to handle and purify the vast majority of peptides.

Hence, the standard strategy for analyzing complex protein mixtures by MS is to digest the proteins into peptides approx 10-20 amino acids long, to analyze these peptide mixtures by mass spectroscopy and finally to infer the composition of the initial protein sample on the basis of the observed peptide mix.

Protein digestion is performed by specific proteases

Proteins are cut into peptides using a well-controlled enzymatic digestion step. The protease most commonly used for this purpose is trypsin, which selectively cuts the peptide backbone of a protein on the C-terminal side of an arginine or lysine residue. Due to the frequency, with which lysines and arginines typically occur in proteins, this leads to peptides with a length of around 10-20 amino acids. Peptides of this size are short enough to be easily handled; yet they are long enough for their sequence not to occur multiple times in a proteome by pure chance.

The sequence specificity of trypsin also is useful to determine a peptides sequence, as the C-terminal amino acid of the generated peptides has to be either an arginine or a lysine.

Peptide sequences are identified by MS/MS

Protein extraction and digestion yields a peptide mixture that may contain millions of different peptides, which need to be separated and identified via two-stage mass spectrometry (MS/MS) (Figure 16).

Peptides are pre-fractionated by HPLC prior to MS

It would indeed be possible to directly apply this complex mixture to the mass spectrometer. Nevertheless, it is advantageous to pre-fractionate the resulting peptide mix before. This is achieved through high-pressure liquid chromatography (HPLC) of the peptide mixture on a reversed phase column (figure 13). In the most common setup, the outlet of the HPLC column is directly connected to the ESI ionization stage of the mass spectrometer. The flow rate of the HPLC column is chosen such that the elution takes place over tens of minutes, which means that the composition of the peptides being eluted from the column changes much more slowly than the subsequent MS/MS analysis step. This gives the mass spectrometer time to analyze the mixture of peptides arriving at the ionization stage in real time.

The first mass filter of the mass spectrometer analyzes the composition of the peptide ions and selects individual peptide species for further analysis

Despite the pre-fractionation step, the sample entering the mass spectrometer will normally still contain multiple different peptides at any given time. The first mass filter analyzes their masses and approximate abundances. Then, the first mass filter is set so that only one of these peptides is permitted to pass to the subsequent analysis, which involves the fragmentation of the peptide into smaller peptides. After the identification of this specific peptide, the mass filter steps to the next one and so forth. Note that many types of protein MS/MS experiments select only a few peptides to be analyzed further. The various strategies for selecting peptides for analysis by fragmentation will be discussed further during the lecture.

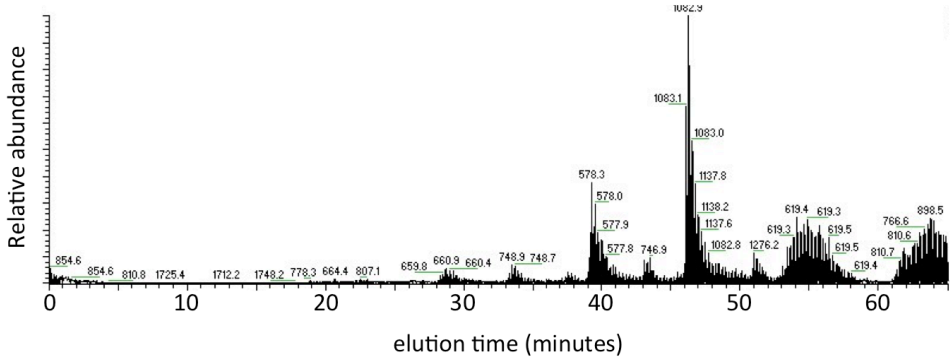


Figure 12 *Elution profile of a complex peptide mixture from an HPLC column. The horizontal axis indicates the time of elution while the vertical axis indicates the amount of peptide being eluted from the column into the mass spectrometer. Prominent peaks in the elution profile are labeled with the mass of the dominant peptide species eluted at that time point*

Controlled collisions with inert gas atoms induces fragmentation of the peptide ions along the peptide backbone

Fragmentation of the peptide is usually achieved by letting the highly accelerated peptide ions collide with inert gas ions in a collision cell. This should not be pictured as a "random explosion" of the peptide into a multitude of random molecular shards. Instead, the conditions in the collision cell are chosen such that a peptide ion only experiences one collision resulting in one bond being broken. This will normally be one of the peptide bonds in the backbone of peptide ions as these are far more susceptible to being broken by one of these collisions than any of the bonds in the amino acid side chains (Figure 13).

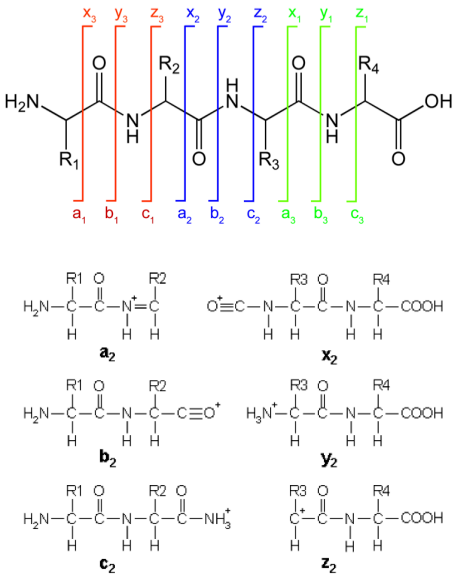


Figure 13 *Fragmentation of a peptide ion following collision with an inert gas is most likely to occur along the peptide backbone. Each peptide unit contains three bonds where the fragmentation can occur and in each case there is a possibility for the ion's charge to either end up on the N-terminal or on the C-terminal fragment. Fragmentation products with their charge on the N-terminal fragment are called a, b and c-ions while those products the charge of which ends up on the C-terminal fragment are called x,y and z-ions respectively. Since neutral fragments stop being accelerated, they are lost from the molecular beam and remain undetected. In practice a,x,c and z-ions are rarely observed. The fragmentation spectra of peptides are dominated by b- and y-ions.*

The breakages occur with comparable, but not identical frequency at each one of the peptide bonds in the peptide. So one might find all possible fragments of a peptide only by breaking a single peptide bond (Figure 14). The mass differences between these fragments correspond to the masses of individual amino acids, which should allow the sequence of the original peptide to be reconstructed. In practice, however, it is rarely possible to directly reconstruct the peptide sequence from the fragment ion sequence. The reverse is much simpler: given a peptide's sequence, its predicted fragment ion spectrum can easily be calculated and compared to the observed spectrum.

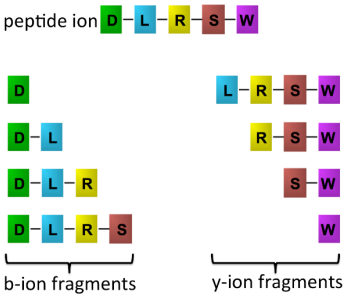


Figure 14 *Schematic representation of the fragmentation ladder generated by the various possible fragmentations of the initial peptide ion. By comparing the mass differences between neighboring peaks in the resulting fragmentation spectra the corresponding amino acid can, in principle, be identified.*

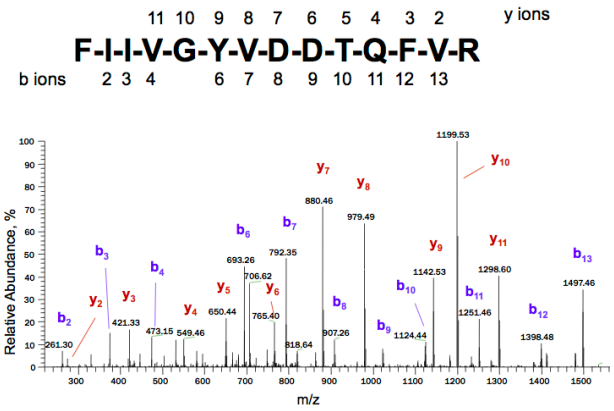


Figure 15 *Fragment ion spectrum of the 14-mer peptide, its sequence indicated in single letter code. The assignment of the peaks to the various fragmentation-generated b- and y-ions is indicated in blue and red letters.*

Observed fragmentation spectra can be matched against predicted spectra to identify peptides

Fragmentation spectra of peptides are rarely as clean and easy to interpret as those shown in Figure 15. Often only a few of the possible b- and y-ions are observed and the spectrum contains a large number of noise. So instead of trying to reconstruct the sequence of the peptide ion from the fragmentation spectrum it is far more advisable to predict the fragmentation spectra of possible peptides and to compare these with the observed fragmentation spectrum. Genomic information of the organism, from which the sample is derived, can be used to predict the sequence of the proteins produced by this organism, in order to simulate the enzymatic digestion of these proteins into peptides and to finally calculate the expected b- and y-ion masses.

With the help of a scoring function the quality of the match between the observed and predicted fragmentation pattern can be assessed. Scoring functions have become increasingly sophisticated and take into account not only the quality of the match between the observed and calculated spectra, but also consider additional information, such as the quality of matches with other candidate peptides or the likelihood for the peptide in question to be observed at a certain stage in the HPLC elution profile. The final result of this analysis is a probability score, which tells us how likely it is for the peptide in question to correspond to a particular predicted peptide.

Identified peptides can be used to infer the presence of their precursor proteins in the original sample.

From the information about the peptides in the sample one can try to infer the presence of a particular protein in the original sample. Again this process involves a series of statistical considerations that have been implemented in specialized software packages. In some cases, a single, highly unique peptide that could be identified clearly via its fragmentation spectrum may even suffice to establish the presence of a particular protein in the sample. But, only if that specific peptide sequence occurs precisely once in the proteome tested. Normally, several peptides need to be identified to make a reliable statement about the presence of a particular protein. It is important to understand that in most cases only a small fraction of the observable ions of a protein are actually seen, but that nevertheless this number is generally sufficient to reliably infer the presence of a given protein. Therefore, modern mass spectrometers can identify thousands of proteins per sample by following this strategy of inference.

Only when the analysis becomes more complicated as in the case of proteins with high sequence identity (e.g. proteins whose genes are derived from recent gene duplications) or in the case of multiple splice variants of a particular protein, it may be difficult to use the peptide-derived information to unequivocally infer the presence of a particular protein in a sample.

Figure 16 gives an overview of what was just discussed on the subject of peptide identification via MS/MS.

Quantification

The preceding section discussed the process of detecting the presence of a particular protein in a sample. In many systems biology applications, however, this information is not sufficient. For example, if one would like to compare the relative abundance of different proteins in a cell before and after a perturbation. Measuring the relative or even the absolute abundance of a protein via MS is not a trivial task.

In MS the intensity of the detector signal correlates poorly with the abundance of the protein in the sample

The intensity of the signal in the detector of a mass spectrometer depends on the efficiency of a number of physical processes, for instance, the efficiency of ionization, the efficiency of fragmentation and finally the

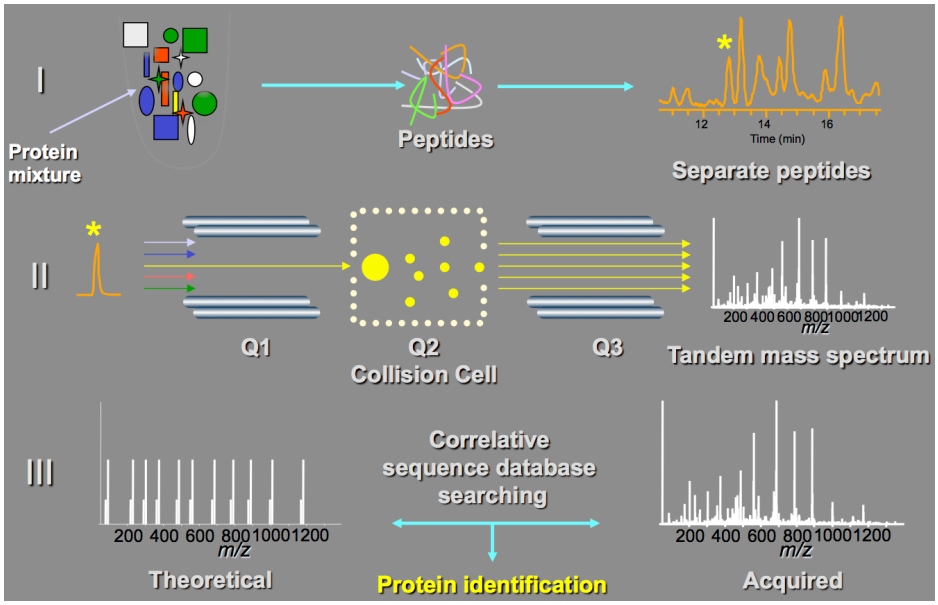


Figure 16 *Procedure for using MS/MS to identify peptides in a complex mixture of protein-derived peptides. It can be separated into three phases:* ***I*** *sample preparation and pre-fractionation of peptides by HPLC.* ***II*** *MS/MS in a mass spectrometer equipped with two mass filters (Q1 & Q3) and an intervening collision-based fragmentation stage (Q2).* ***III*** *identification of individual peptides by matching their observed fragmentation ion spectrum to a database of all theoretically possible fragmentation spectra predicted from the sequence of the organism's protein coding genes.*

efficiency with which an ion triggers a signal in the detector. The intensity also depends on the type and

abundance of other molecules present during ionization because the different molecules compete for a limited number of charges in the electrospray droplets. This latter effect is called ion suppression. These individual efficiencies vary greatly across different peptides, so that the overall efficiency of signal generation may vary across peptides by several orders of magnitude. In other words, the intensity of the detector signal generated by a peptide ion is a very poor indicator for its abundance in the sample.

Even for the same peptide the efficiency in generating a signal will fluctuate from day to day and sample to sample due to subtle drifts in instrument settings and interactions between the peptide and sample matrix. So even when comparing signal intensities of the same peptide from two different samples, these relative intensities are only a rough indicator of the peptide's relative abundance in the two samples.

Complex sample preparation steps hamper reproducibility of results

As we have seen above, the path from the original biological sample to the recording of the mass-spectrum comprises many different steps. Further, these sometimes difficult to reproduce exactly in successive cycles, thus further reducing the reproducibility of MS-based protein abundance measurements across multiple samples.

Selective isotope labeling of samples allows the simultaneous measurement of multiple samples and enables reliable comparison of protein abundance across samples.

Currently, all reliable strategies for comparing the abundance of proteins across samples via MS are based on stable isotope labeling. As discussed above, isotopes have identical chemical properties, but different masses and modern mass spectrometers are able to detect these mass differences reliably. Isotope-based labeling makes use of this property to separately track chemically identical peptides from different sources in a single pooled sample. The general principle of isotope labeling is explained here with the example of SILAC (stable isotope labeling by amino acids in cell culture) (Figure 17), but variations of this specific strategy can be applied to make quantitative measurements in a variety of experimental settings (Figure 18).

A standard systems biology experiment is the perturbation of a cell line (either by mutation, a change in growth conditions or the addition of a medicinal drug), followed by an analysis of the way in which the perturbation changed the state of the cell (e.g. change of the relative abundance of different proteins). In a SILAC experiment two cell cultures are grown in parallel. The first cell culture, which represents the unperturbed condition, is grown in standard growth medium. The second culture is grown on a medium that although being chemically identical to the growth medium of the first culture, includes

amino acid variants containing only heavy isotopes. This second culture is exposed to a perturbation which could induce a change in the cell's protein composition. Subsequently, both cultures are pooled and undergo protein extraction, protease digestion and mass spectrometry as a single sample. In the mass spectrum, however, two different signals are readily distinguishable. For due to their differing mass, the ions originating from the perturbed sample give rise to a signal that is clearly distinct from the one that the ions from the unperturbed sample produce. This enables, an easy and accurate comparison of the relative abundance of equivalent peptides in two different samples.

The same fundamental principle of isotope labeling is used in many variations (Figure 18) in order to determine the relative and even absolute abundance of proteins and peptides. Isotope labels are either introduced by incorporating isotope labeled amino acids during natural protein synthesis or are chemically introduced after the cells are disrupted. Alternatively, the sample can be spiked with isotope-labeled versions of specific peptides or proteins. These reference peptides can be either synthesized chemically or by recombinant expression in microorganisms grown on isotope labeled growth media. Although the fact that a specific reference peptide has to be generated for each of the peptides to be quantified is clearly a disadvantage, it should not be forgotten that the spiking strategy also allows the amount of labeled peptide being added to be measured, thus enabling the determination of absolute concentrations.

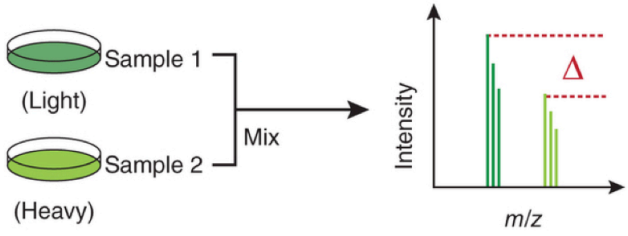
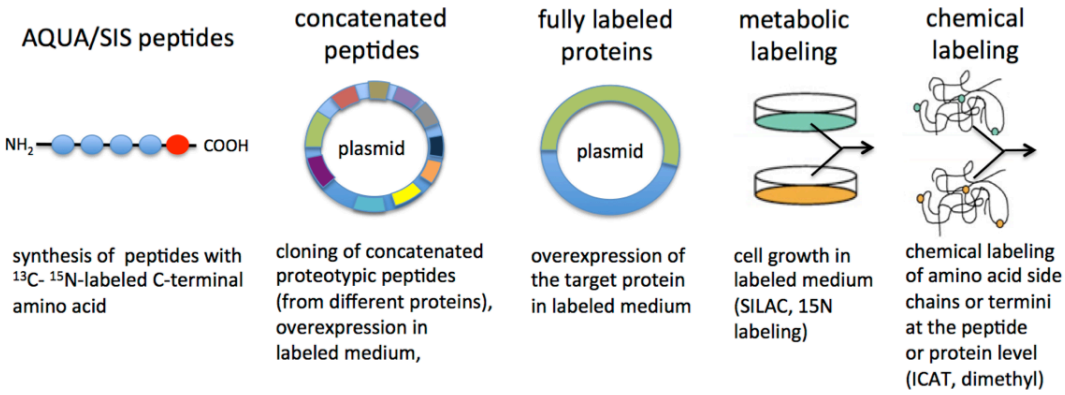


Figure 17 *SILAC experiments allow the measurement of changes in the relative abundance of proteins in response to a perturbation of the system. Two identical cell lines are grown in culture on chemically identical growth media. In the growth medium of one of the cells cultures, however, the amino acids are replaced by their heavy isotope variants. One of the two cell cultures is then perturbed. The cells from the two cultures are pooled and processed together. In the mass spectrum the peptide ions originating from the two cultures can nevertheless be distinguished due to their difference in mass, thus enabling the accurate detection of changes in the relative abundance of proteins between the two samples.*



relative quantification	+	+	+	+	+
absolute quantification	+	+	+	-	-

Figure 18 *Comparison of different isotope labeling strategies for the quantification of peptides and proteins in complex samples via mass spectrometry*