Recitation 11 Introduction to MS methods **April 20-22 2016**

Required reading: **Seo/Carroll and Angew Chem Int Ed 2011 50 134-45 and SI.** Paulsen et al Nat Chem Biol 8 57-64 and SI (2012) covered in recitation 12 using the methods in recitation 11.

I. Mass spectrometry of biomolecules

Protein analyte → "Soft ionization (charged ions in gas phase without destruction"

Mass spectrometry (MS) measures the mass to charge ratios (m/z) of gas-phase ions. The most commonly used MS of biomolecules utilizes so-called "soft" ionization methods: **electrospray ionization (ESI-MS)** and **MALDI-TOF MS (matrix-assisted laser desorption/time of flight MS)**. Large biomolecules, such as proteins, are polar and thermally labile, and so are not easily vaporized and ionized by conventional MS methods without significant fragmentation. Note that ESI and MALDI utilize proton ionization, generating $[M+nH]^{n+}$ or $[M-nH]^{n-}$ ions. Traditional MS methods utilize electron ionization, generating $[M\cdot]^{+}$ species that can easily fragment.

I. Electrospray ionization MS

1. Method:

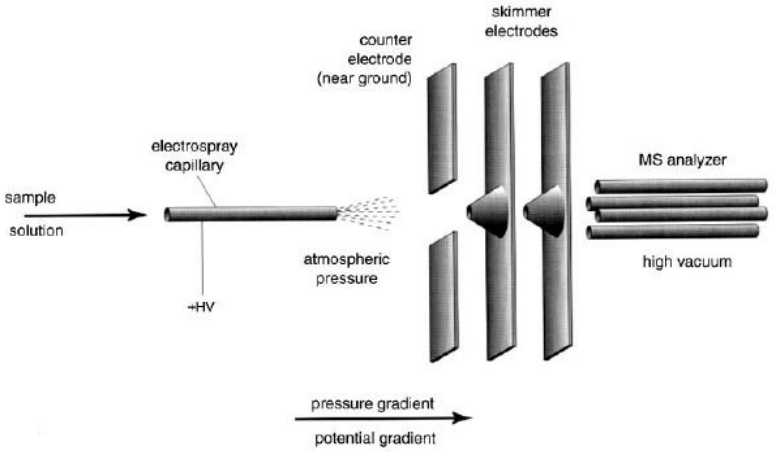


Figure 1. Essential features of the electrospray interface.

In ESI-MS, a solution containing the molecules of interest (in some ionized form) is pumped through a thin capillary. At the tip, small charged droplets (each containing *many* molecules of analytes) are sprayed into a chamber at atmospheric pressure and travel down a potential and pressure gradient to an analyzer.

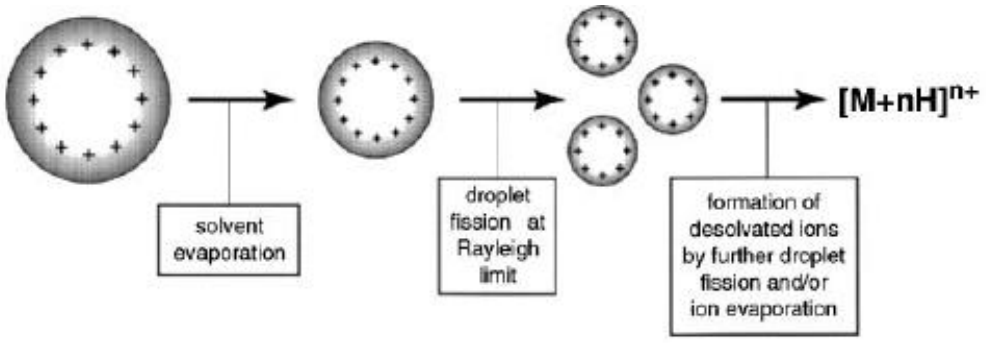
S.J. Gaskell, *J. Mass Spectrom.* 32, 677, 1997

Large droplets with many charged analytes become isolated, desolvated ions through a combination of two mechanisms:

1. Solvent evaporation (sometimes assisted by application of a heated gas) from small droplets increases ion concentration.
2. As charge density within a droplet increases, it reaches a point where surface tension can no longer sustain the charge anymore (the

Rayleigh limit). The droplets fragment into smaller droplets via a "Coulombic" explosion.

These two mechanisms are iterated until single gaseous ions are produced.

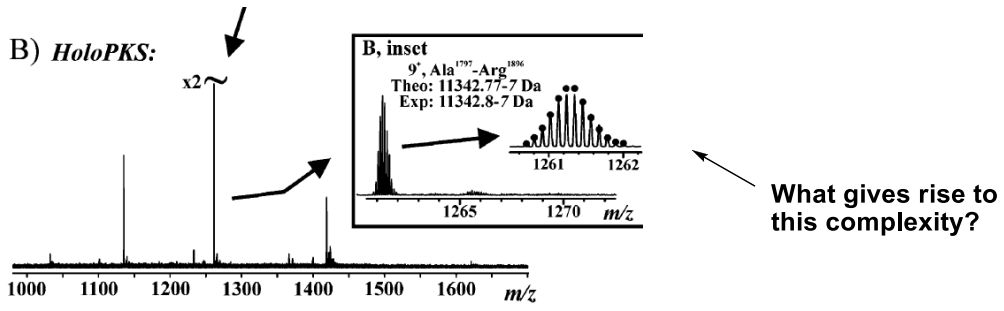


Samples are usually prepared in an acidic solution to generate $[M+nH]^{n+}$ ions.
ESI-MS is well suited for high-resolution analysis of a range of biomolecular masses. The fact that proteins have many sites that can become positively charged means that even very high molecular mass proteins will be amenable to detection by a mass analyzer with a m/z range of 1000-2000 Da.

2. Mass envelopes:
Assume that we have a protein of 10000 Da. The following m/z signals might be detected:

Ionization state	m	m/z
+1	10001	10001
+2	10002	5001
+3	10003	3334.3
+4	10004	2501
+5	10005	2001
+6	10006	1667.7
+7	10007	1429.6
+8	10008	1251

This gives us a mass envelope. Below is an example of a sample for analysis of a domain of a PKS.

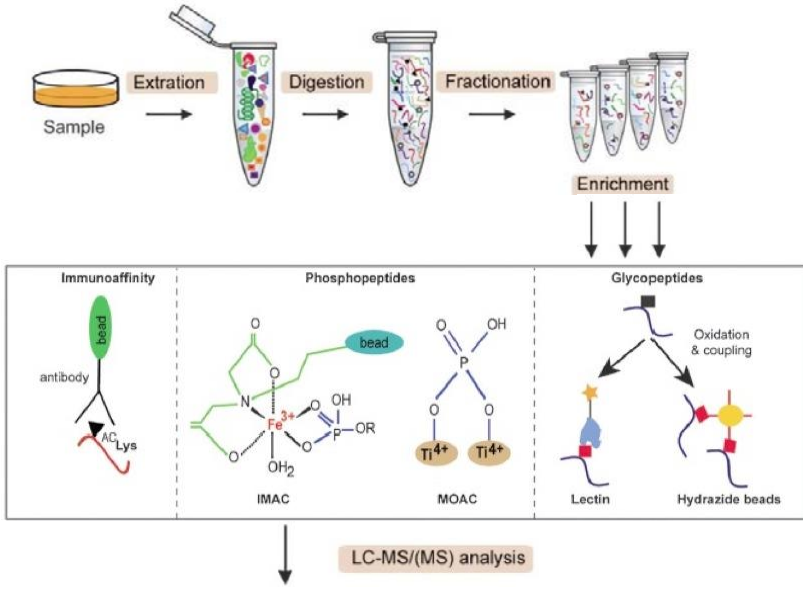


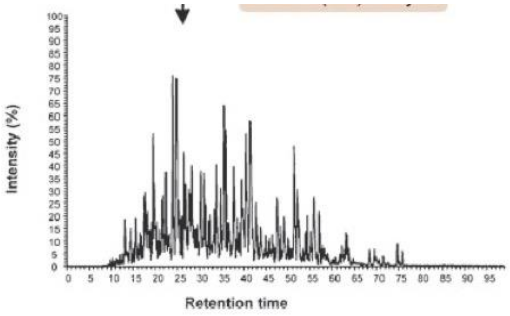
Since soft ionization methods like ESI and MALDI are designed not to fragment the parent ion, a collision induced dissociation step is added after the first MS step (tandem MS). This method usually involves the selection of a certain parent ion for further analysis. An inert gas is introduced into a collision chamber, and collisions between the inert gas and the selected ion lead to its decomposition into several product ions, which are then mass-analyzed.

II. Sample preparation for proteolyzed fragments separated by Reverse phase HPLC (LC/MS/MS):

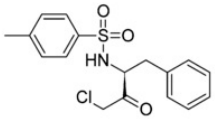
Often to simplify analysis, proteins are digested into peptide fragments using a protease (trypsin, pepsin in the EFGR paper that will be discussed in recitation 12) before MS analysis. The peptide fragments are amenable to high-resolution detection and MS/MS sequencing. For the proteolysis step, trypsin, a serine protease, is often use as it reliably cleaves adjacent to Lys and Arg residues. The source of trypsin is important as it is often contaminated with other serine proteases such as chymotrypsin. Thus in the experimental sections of many papers you will see that trypsin has been treated with the protease inhibitor TPCK (tosyl phenylalanyl chloromethyl-ketone shown below). This inhibitor inactivates chymotrypsin that would add complexity to the predicted peptide fragments.

Reversed phase HPLC (High performance liquid chromatography) is also commonly used to separate the resulting peptides to simplify the MS analysis. The designation "reversed phase" refers to the stationary phase being nonpolar (e.g. large hydrocarbons, commonly C-18 or C-8) and the mobile phase is polar. Peptides are separated by their relative affinities for the stationary phase when a buffer of increasing hydrophobicity is applied. Below is a typical separation of a proteolytically cleaved protein.



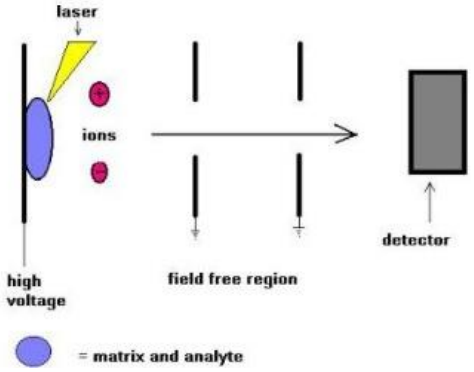


Reverse phase HPLC trace (monitored at 220 nm)



TPCK

II. MALDI-TOF MS (matrix assisted laser desorption coupled to time of flight due to large mass range)



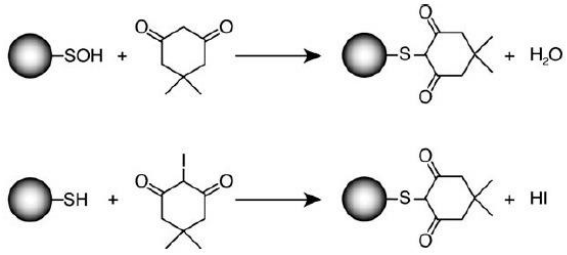
In the first step, the protein of interest is mixed with a matrix in solution and put on a metal plate. The solvent evaporates leaving a co-crystallized matrix/analyte. The matrix material absorbs the UV laser light and primarily the matrix is desorbed and ionized by this event. The top 1 micron is converted into a hot plume of matrix clusters and nanodroplets and the analyte molecules are ionized (protonated or non-protonated) in this plume. The matrix is then thought to transfer a proton to the analyte thus charging them. Ions observed consist of a neutral molecule [M] and an added or removed ion $[M+H]^+$, [M+Na]+ or $[M-H]^-$. The matrix consists of crystallized molecules such as sinapinic acid or 3,5-dihydroxybenzoate. Identification of the best matrix is determined by trial and error, but all matrices are low molecular weight and easy to vaporize. They are often acidic providing a proton source. They have a strong optical absorption with a chromophore in the UV.

III. Focus on technology development in Angew Chem Int Ed 2011 50 134-45

In the papers that you have read and will read (recitation 12), we are concerned with development of technology for identification and quantification of post-translational modifications (PTMs) generated inside a cell that result in a mixture of modified and unmodified proteins. Specifically, we are focused on the issue of sulfenylation which is a reversible modification thought to play an important role in signal transduction mechanisms. For technology development we will focus on a model protein glutathione peroxidase 3 (Grx3) and for analysis of PTM, the focus will be on the epidermal growth factor receptor (EGFR), a tyrosine kinase, whose

activity can be modulated by sulfenylation. The problem inside cells is a challenging one as there are a huge number of proteins in cells, there are many different modifications, and the concentrations of the proteins can span 12 orders of magnitude.

1. Reagent to detect sulfenylated cysteine: Dimedone (Dm) and iodo-dimedone (I-Dm) derivatives were developed to detect sulfenic acid and unmovided thiols, respectively in the protein of interest (POI).



Sulfenylation (CH_2-SOH) occurs by reaction of a cysteine with H_2O_2. Over oxidation to CH_2-SO_2^- and CH_2-SO_3^- can occur. The last modification is irreversible. These modifications will be briefly discussed in Module 7.

2. **The method of detection of PTM of modified cysteines by MS:** To test the methodology, a well characterized protein, Gpx3, was chosen and examined. Gpx3, (22,736. 8Da), has two cysteines that were mutagenized to serines (C64S, C82S) leaving a single cysteine (C36). This protein treated with I-Dm (Dm is 138 Da), gives a protein of 22874.6 Da, consistent with a single modifcation. Below is the ESI-MS of the unlabeled and labeled species (a, left). Note that both Dm and I-Dm give the same modified protein, but the former only reacts with sulfenic acids and the latter with thiols. Deuterium labeling experiments described in the required reading, offer a way to distinguish between the two. For these reagents (Dm and I-Dm) or any reagent) to be useful, the reactions need to occur quantitatively in aqueous solutions under mild conditions at neutral pH.

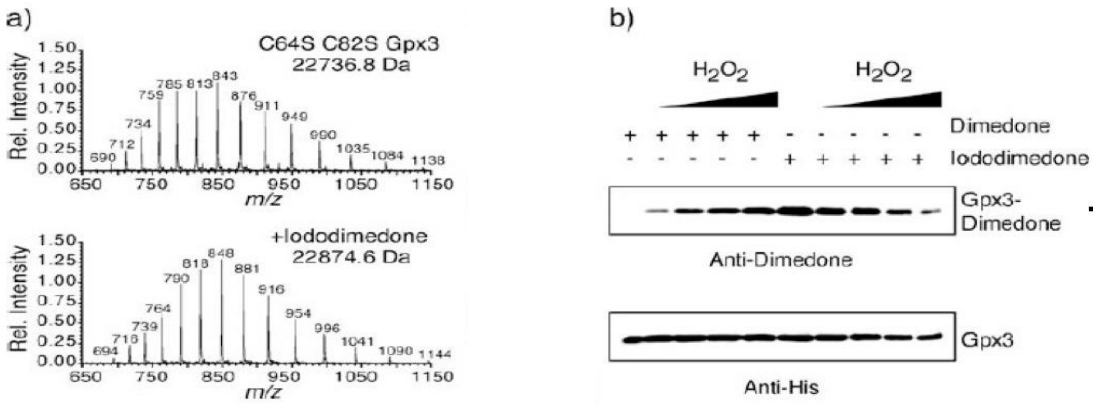
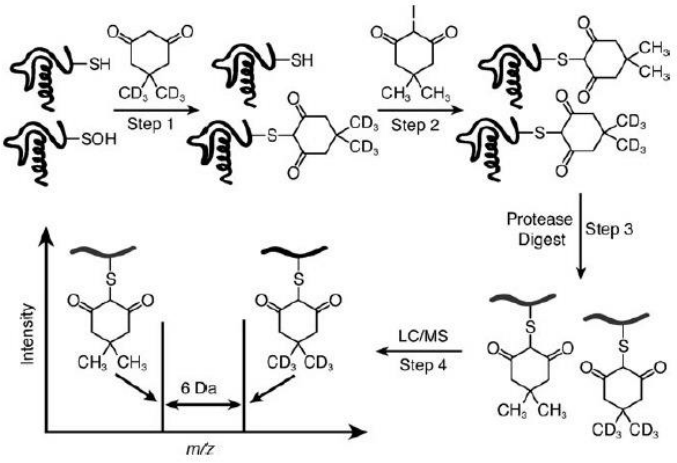


Figure 1. ESI-LC/MS and Western blot analysis using Abs to Dm shows that iododimedone selectively labels the thiol group of C36 in C64S C82S Gpx3. a) The molecular weight of mock-treated C64S C82S Gpx3 is 22736.8 Da corresponding to intact, unmodified protein (top spectrum, left). The molecular weight of C64S C82S Gpx3 incubated with iododimedone is 22874.6 Da (bottom spectrum, left), corresponding to C64S C82S Gpx3 with a single dimedone adduct (Dm=138 Da). b) C64S C82S Gpx3 (50 μM) untreated or oxidized with H_2O_2 (12.5, 25, 50, or 100 μM) followed by Dm (20 mM) or I-Dm (20 mM). The resulting

samples were analyzed by Western blot using an antibody that recognizes the protein–dimedone adduct (top panel, right). Total protein content was determined by an anti-His antibody (bottom panel, right). Panel b shows with Gpx3 that with Dm as you increase the amount of hydrogen peroxide, you increase the amount detected with the Dm Ab and with I-Dm you decrease the amount of free cysteine.

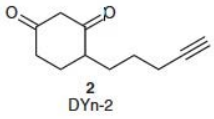
3. Method to quantitatively distinguish between the two types of cysteine modification. An example is shown below. [D_6]-Dm and unlabeled I-Dm are used as "D" minimally perturbs the reagent with respect to structure and chemical reactivity. [Question could I-Dm react with other



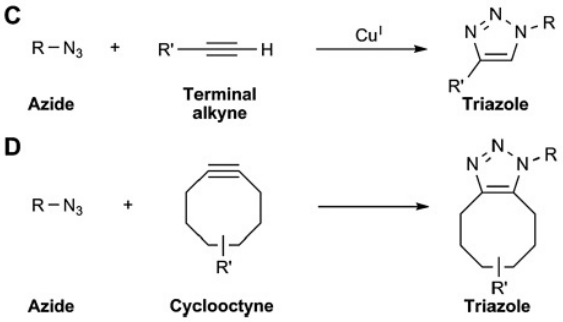
amino acid side chains such as lysine or imidazole?] The POI is then first labeled with [D_6]-Dm, then the excess reagent is removed, and finally the sample is treated with unlabeled I-Dm. The protein is proteolyzed and the peptides separated and analyzed by LC/MS (liquid chromatogapy/MS). Quantitation is assessed by measuring the **ratio** of the heavy peak to the heavy + light peak (see procedure above).

As a control experiment the authors used glyceraldehyde 3-P dehydrogenase (GAPDH) that has a cysteine that is essential for catalysis. They incubated GAPDH with hydrogen peroxide and the Dm and I-Dm reagents. In this case, the experiment failed because the sulfenic acid was further oxidized to sulfinic and sulfonic acids, before the sulfenylation was complete. Dm does NOT react with these modified cysteines. Thus the ratio of sulfenylated protein to unmodified cysteine was not reciprocal. The reactivity of the cysteine with the hydrogen peroxide in the particular protein being investigated, thus plays plays a very important role in the outcome (data in paper).

IV. Dm reagent for use in cells (from the Nat Chem Biol paper as an introduction): For these reagents (Dm or I-Dm) to be useful inside the cell, they need to have a "handle" (an acetylene or an azide, see cartoon below) to allow affinity purification of the protein of interest (POI), that is, to remove the POI away from the cellular mess to enhance sensitivity. In the Nat Chem Biol paper, the reagent found to be most useful was DYn-2. This reagent is membrane permeable so that it gets into cells. Of course you also need to worry about the length and bulk of the labeling arm (they tried a number of reagents in the paper) and whether this linker still allows the reagent to get into the active site to label the modified cysteine.

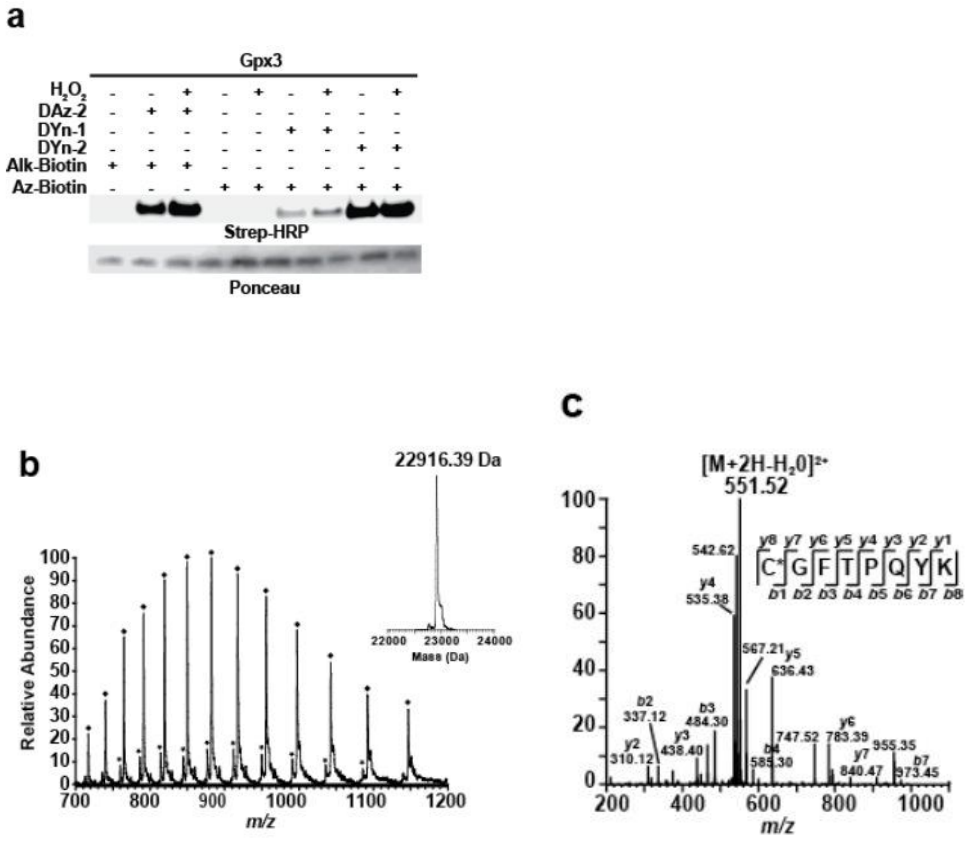


The acetylene can react with a biotin-linker-azide via click chemistry to give a biotin tagged-POI that can then be purified by affinity chromatography with streptavidin. Alternatively, Carroll et al made dimedone derivatives with azide linkers, which can react with a biotin-linker-alkyne. The chemistry between the azide and acetylene, can be copper catalyzed. In the absence of copper, not useful inside the cell as it is toxic, a strained alkyne such as cycloctyne may be used.



For a review on these methods see: Shieh and Bertozzi (2014) Org and Biomol Chem 12, 9307.

The data presented below is from SI Figure 4 from the Carroll Nat Chem Biol paper. It again illustrates proof of concept using Gpx3 to assess modification with different dimedone reagents for use inside cells.



SI Figure 4 from Nat Chem Biol paper: a. Comparison of DAz-2, DYn-1 and DYn-2 (reagents synthesized by Carroll et al) for detection of sulfenic acids using Grx3 as a model, for use in vivo. b. ESI-LCMS intact mass analysis shows the covalent attachment of a single DYn-2 molecule Gpx3 (22916.39 Da). The inset shows the deconvoluted mass spectrum. The observed mass shift is 177.48 Da and the expected mass shift for a single DYn-2 adduct is 176.08 Da. C. ESI-LC/MS/MS analysis of the reaction using collison-induced dissociation (CID) MS of the precursor ion m/z 551.52 $[M + 2H- H_2O]^{2+}$ corresponding to Dyn-2-tagged peptide (CGFTPQYK-OH) derived from a trypsin digest and analysis of each of the peptide fragments.

5.08J Biological Chemistry II

Spring 2016