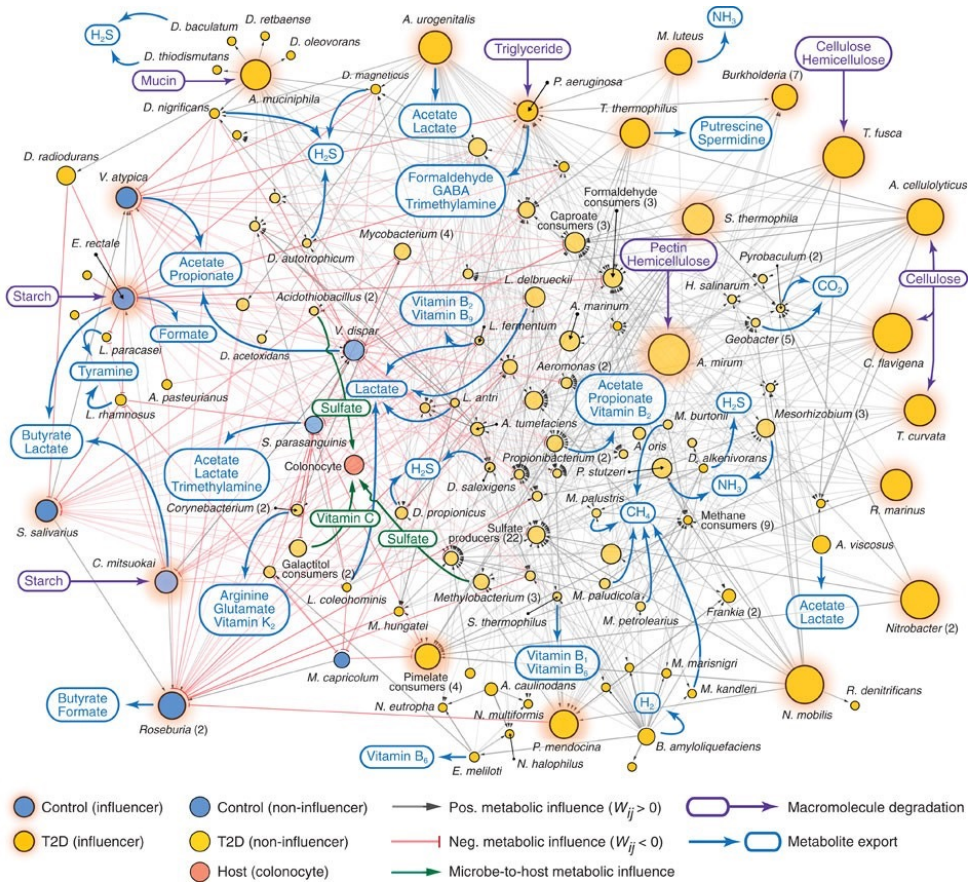
Intro to mass spectrometry

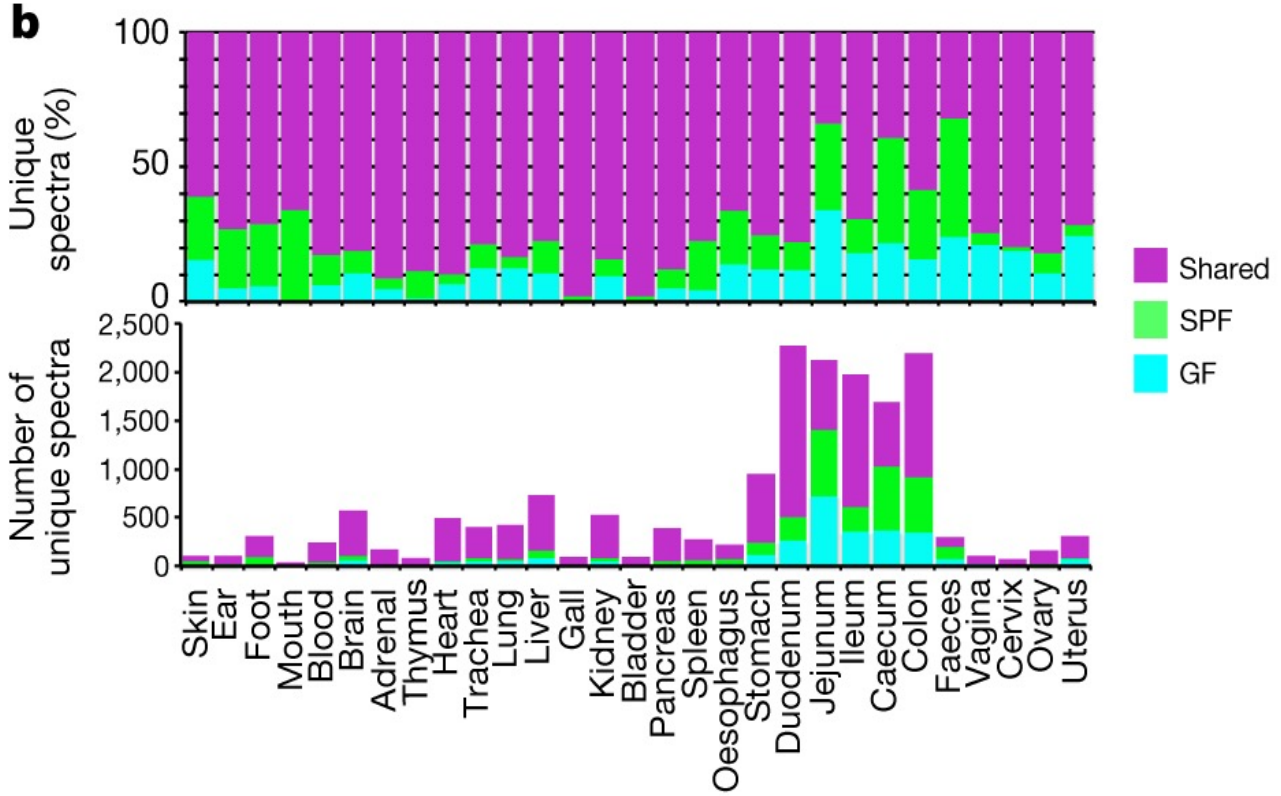
- Objectives
  - Apply knowledge of mass defect, exact mass, and resolution to interpreting mass spec data

  - Intro to metabolomics (mass spec on mixtures)

Moriah Sandy, QMAC director
Moriah.sandy@ucsf.edu

Why is mass spectrometry important for studying the human microbiome?

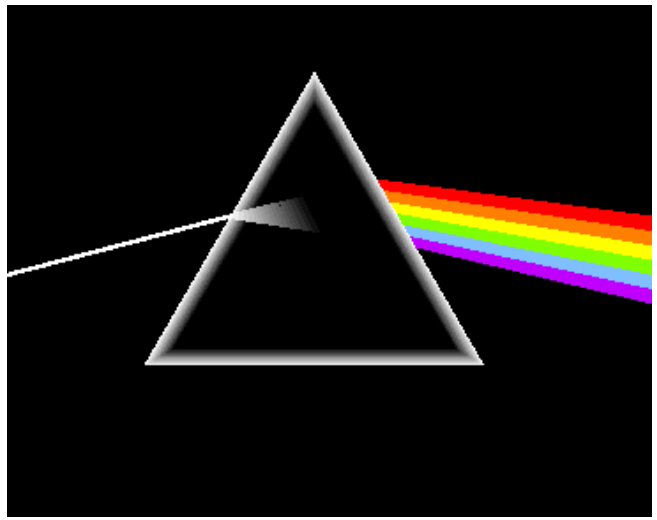




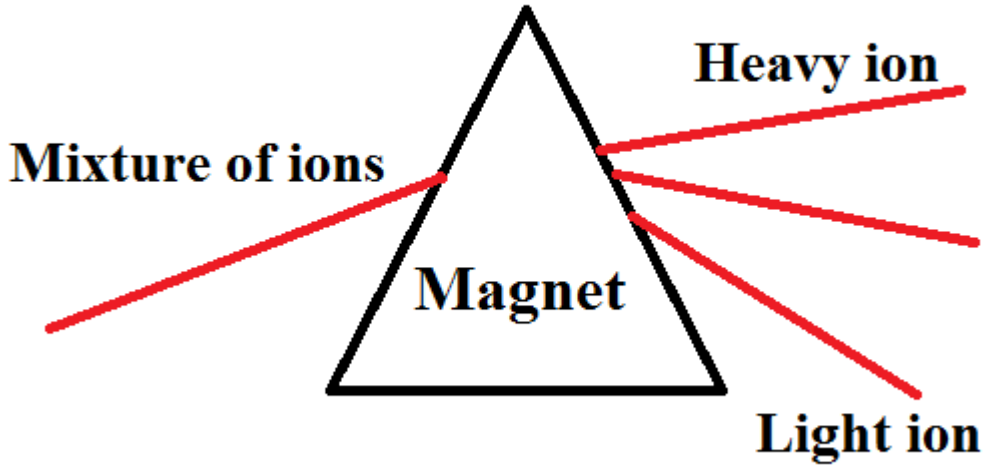
Sung *Nature Communications* 2017
Quinn *Nature* 2020

Fundamentals

- Can only detect IONS not neutral molecules
- First step in any MS analysis is to convert a neutral molecule into a charged ION
- MS separates ions by mass to charge ratio (*m/z*)

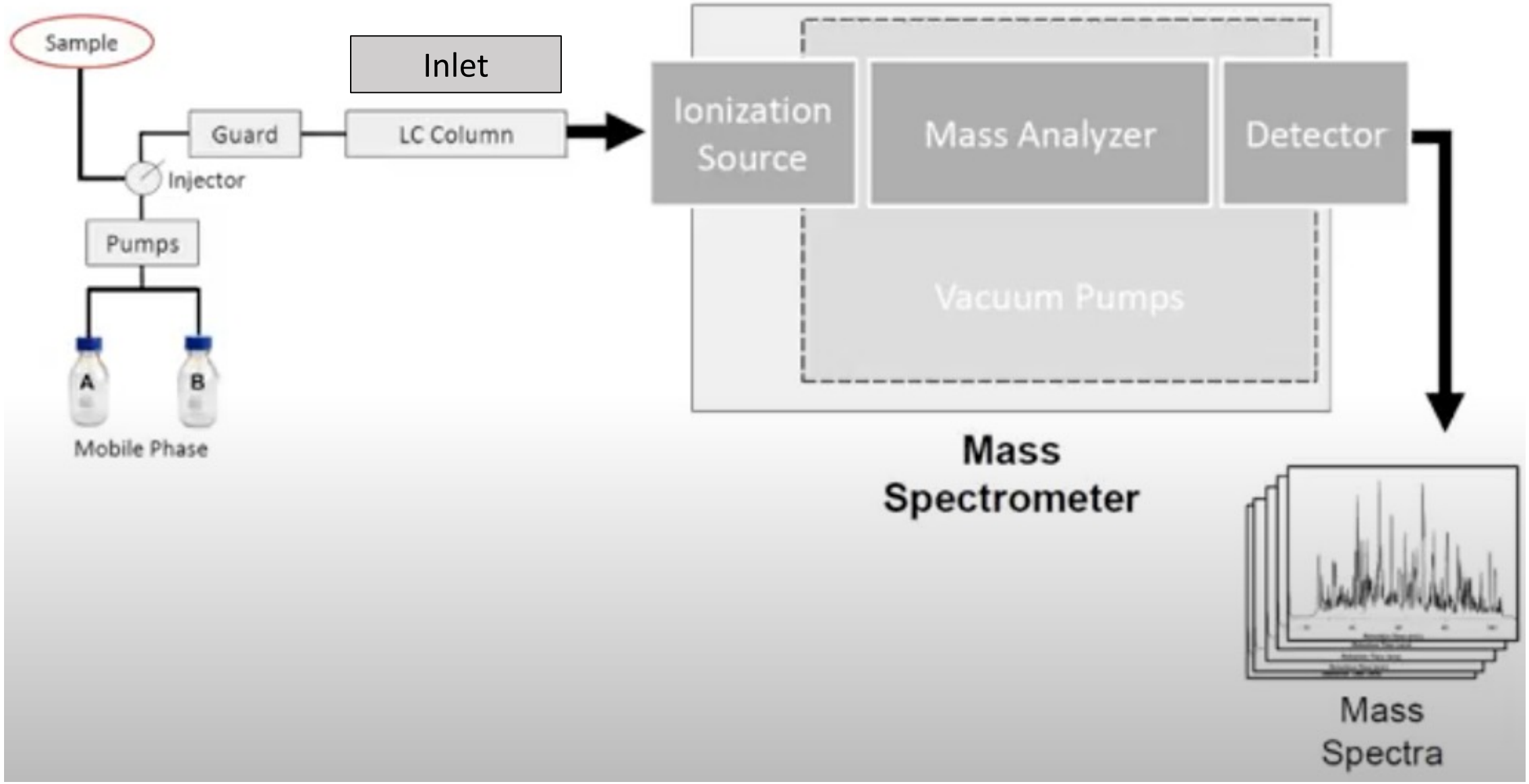


Prism (light spectrometer) deflects blue light more than red light

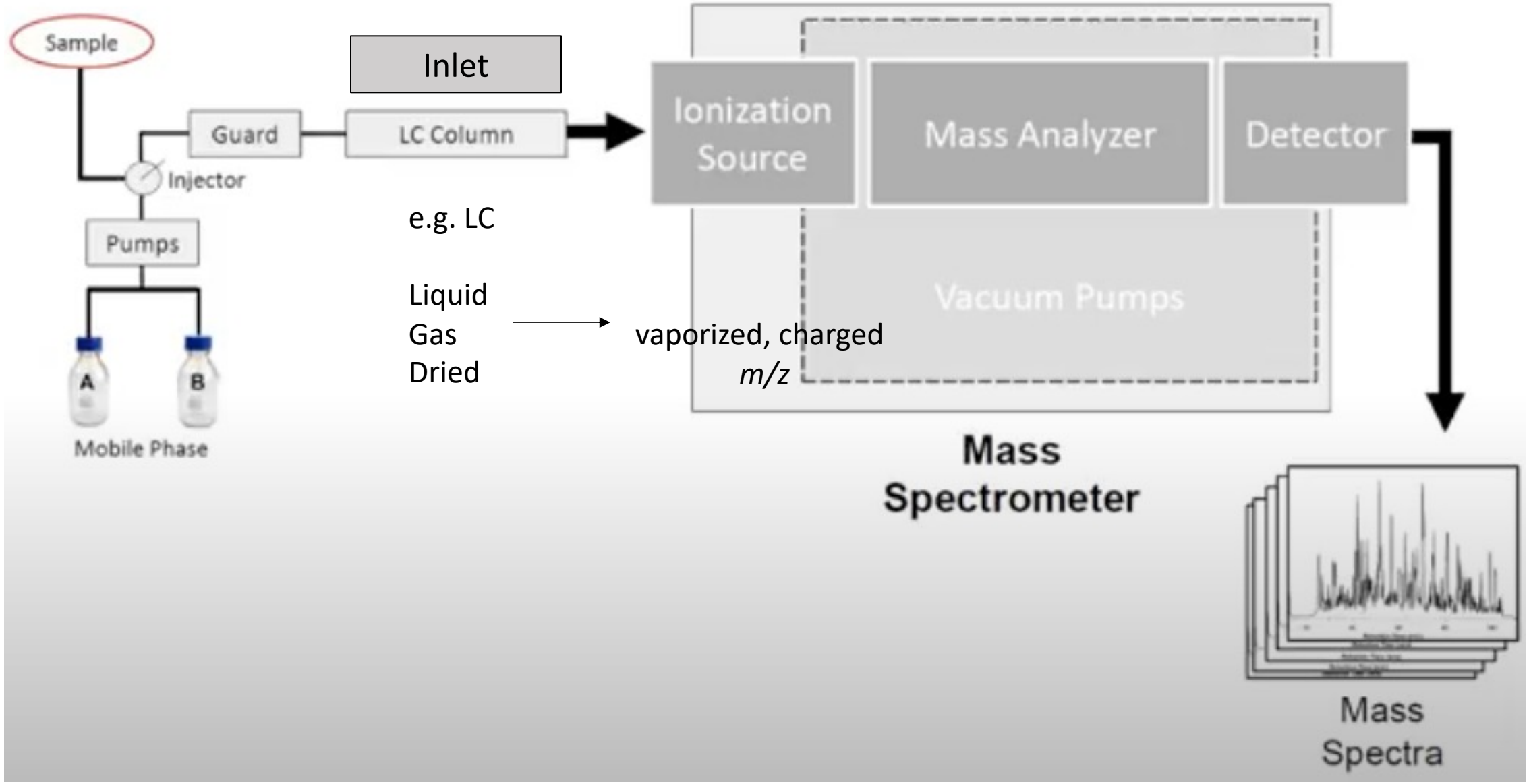


Magnet (mass spectrometer) deflects light ions more than heavy ions

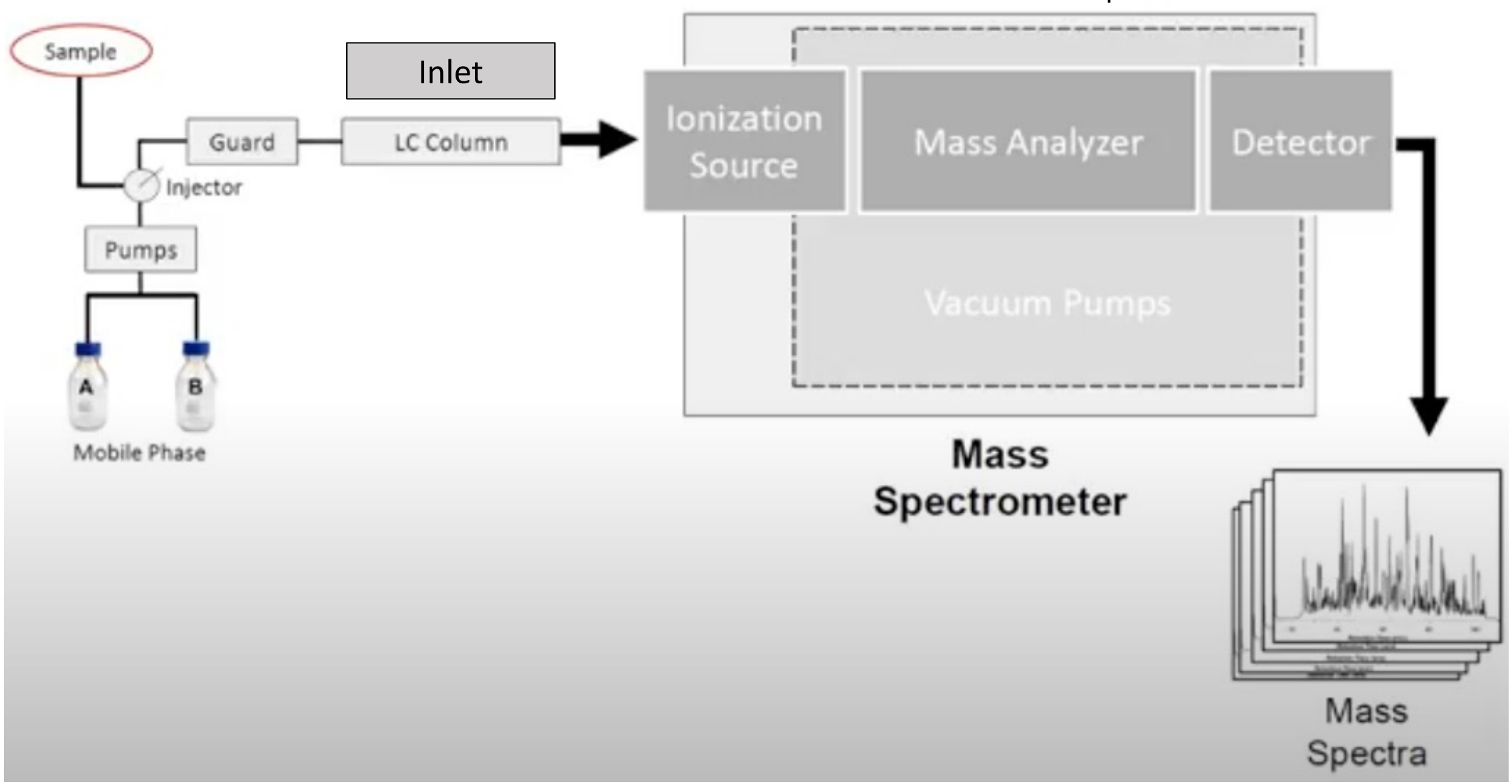
Instrument Basics



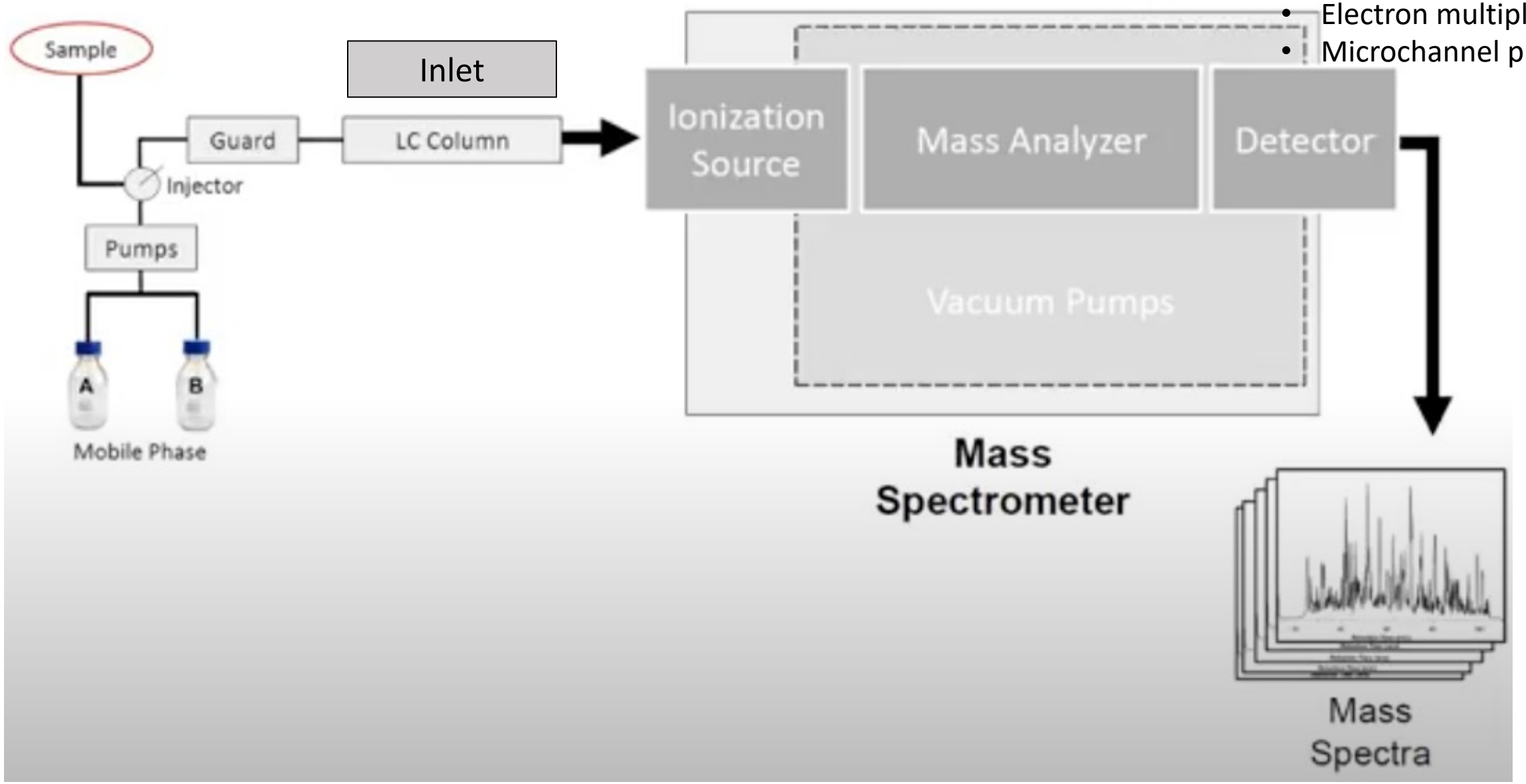
Instrument Basics



Instrument Basics

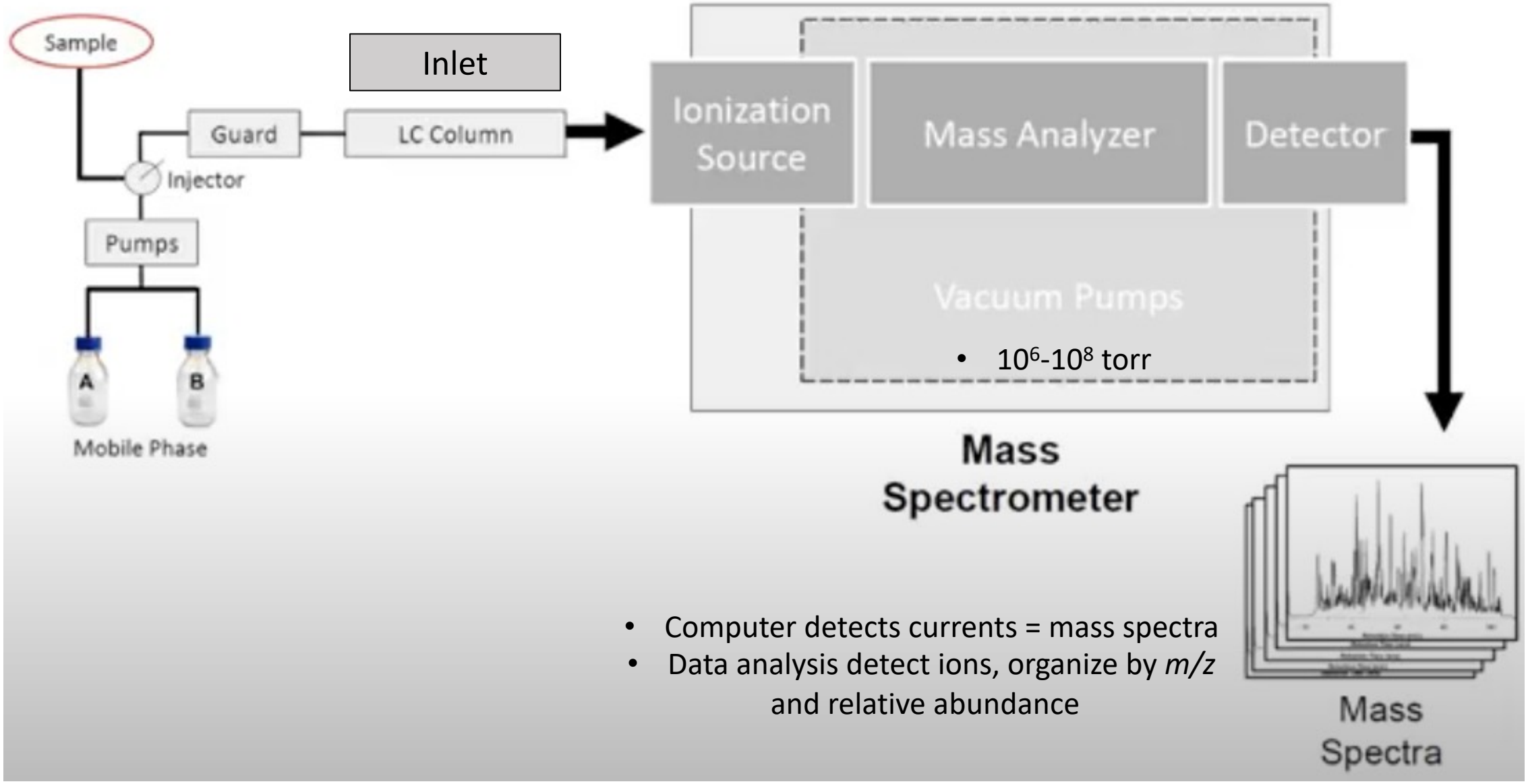


Instrument Basics



- Electron multiplier
- Microchannel plate

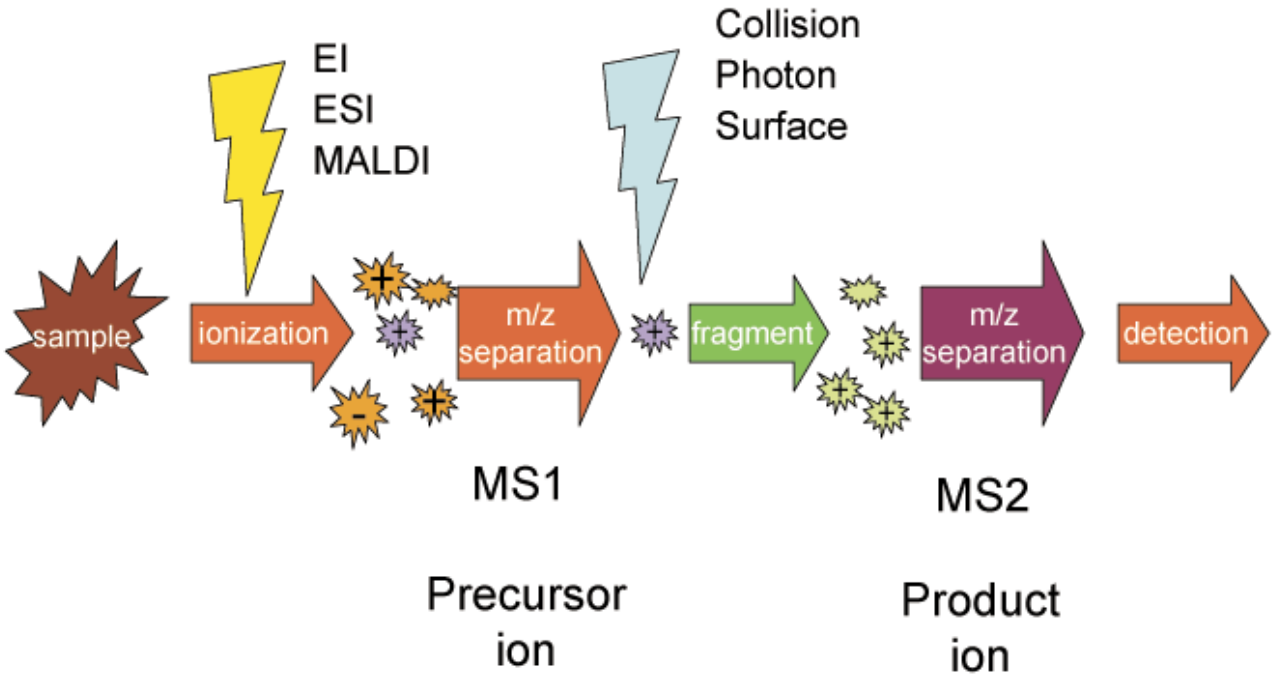
Instrument Basics



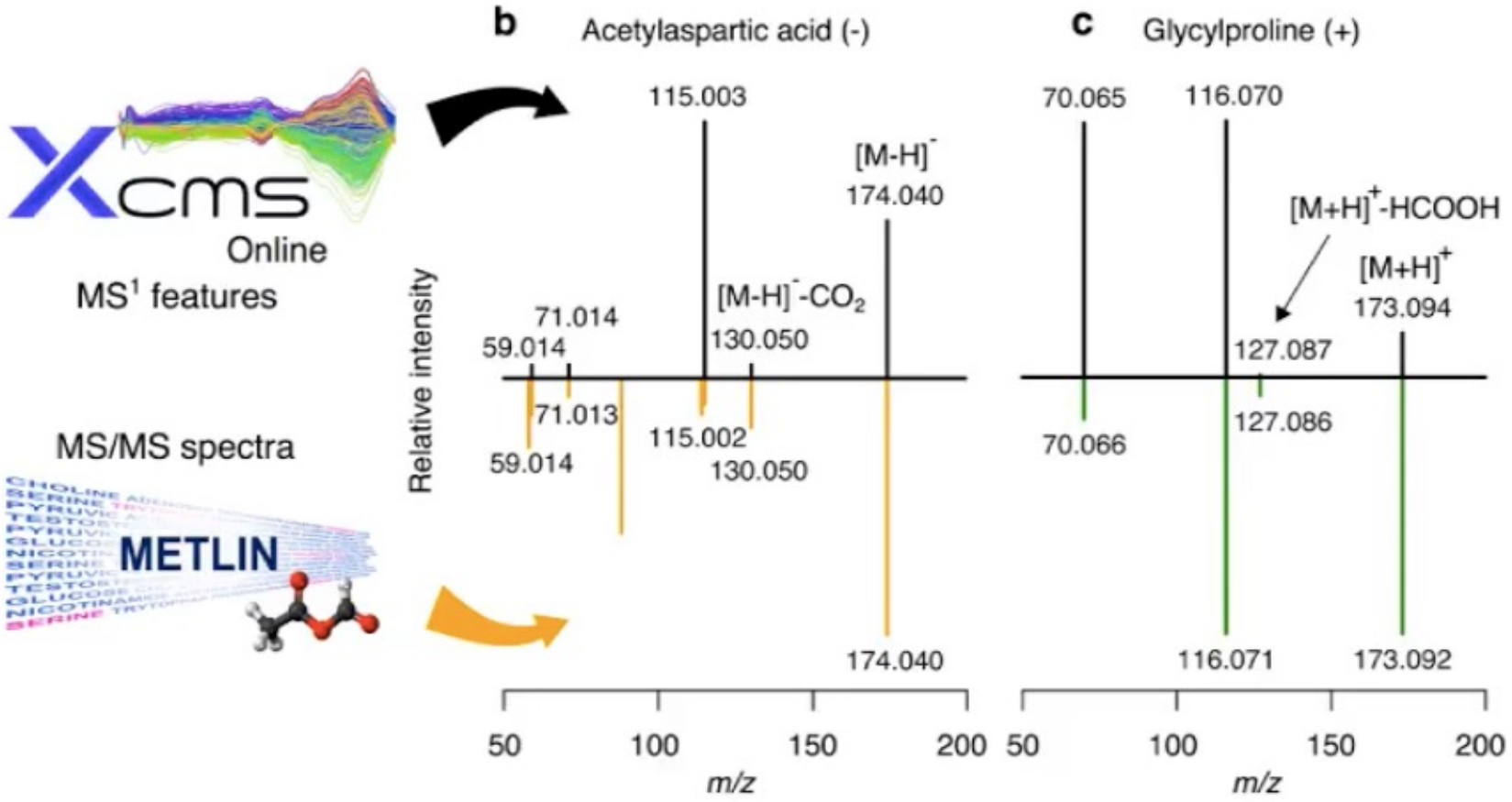
- Computer detects currents = mass spectra
- Data analysis detect ions, organize by *m/z* and relative abundance

Tandem Mass Spectrometry

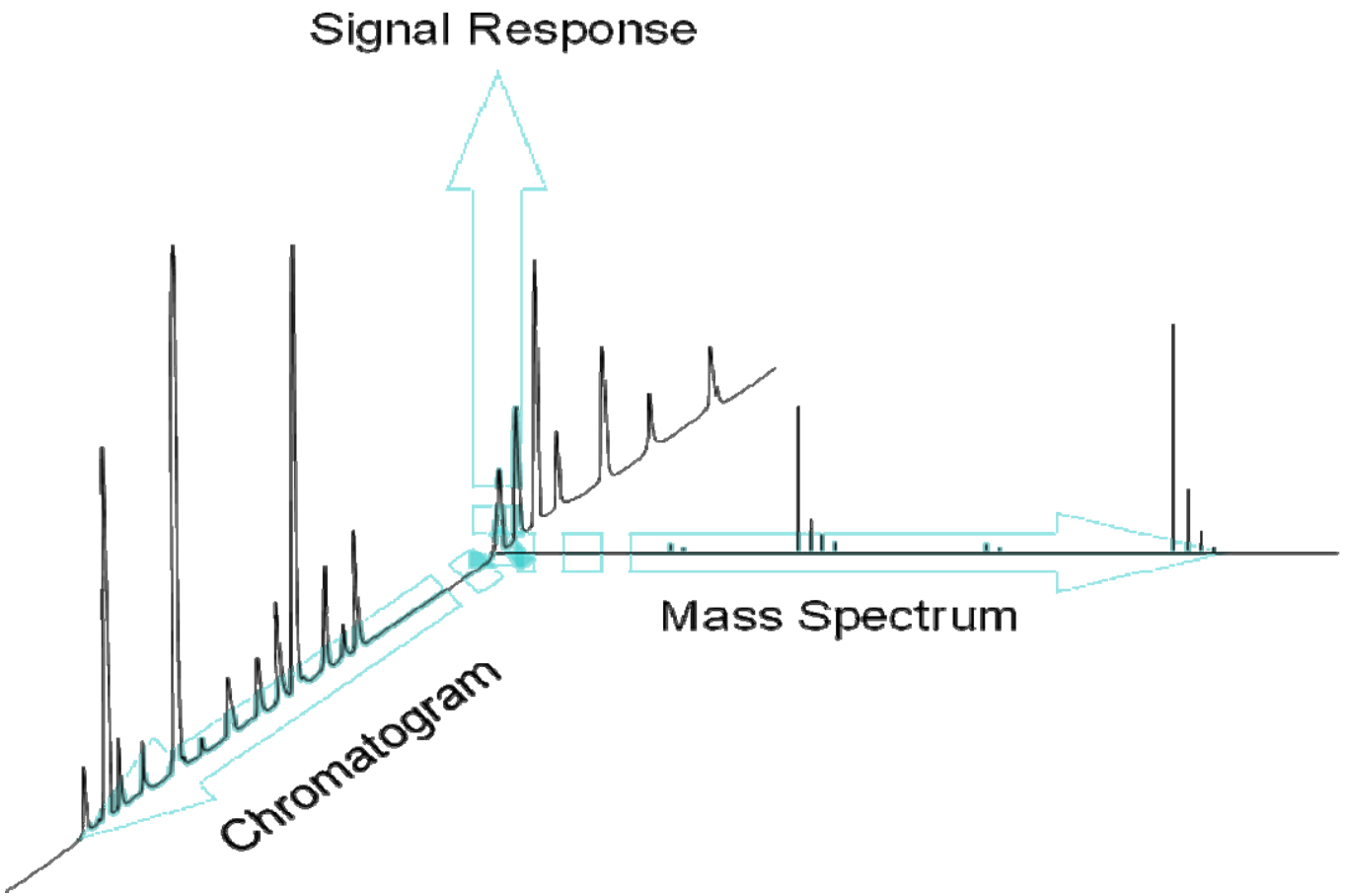
- Tandem MS (aka MS/MS)typically involves the use of 2 or more mass analyzers (exception ion trap).
- MS1 Ion *(m/z)* selection + fragmentation usually by collision with inert gas (e.g. Argon)
- The product fragment ions (MS2) are then detected and analyzed.
- Quadrupole, Ion Trap



Tandem Mass Spectrometry



Chromatography + Mass Spectrometry



M. Sargent (Ed.), Guide to achieving reliable quantitative LC-MS measurements, RSC Analytical Methods Committee, 2013.
ISBN 978-0-948926-27-3.

Atomic Mass Scale

1 atomic mass unit (1 amu, 1 Da , 1 *u*) is 1/12 of the mass of carbon $_{12}^{6}C$

i.e. 6 protons + 6 neutrons + 6 electrons = 12 amu

BUT! If we look up the masses of protons, neutrons, and electrons we get:

Proton = 1.0073 Electron = 0.0006 Neutron = 1.0087

(6 x 1.0073) + (6 x 0.0006) + (6 x 1.0087) = 12.0996 amu

Mass Defect = (Actual Mass – Nominal Mass) = 0.0996 amu

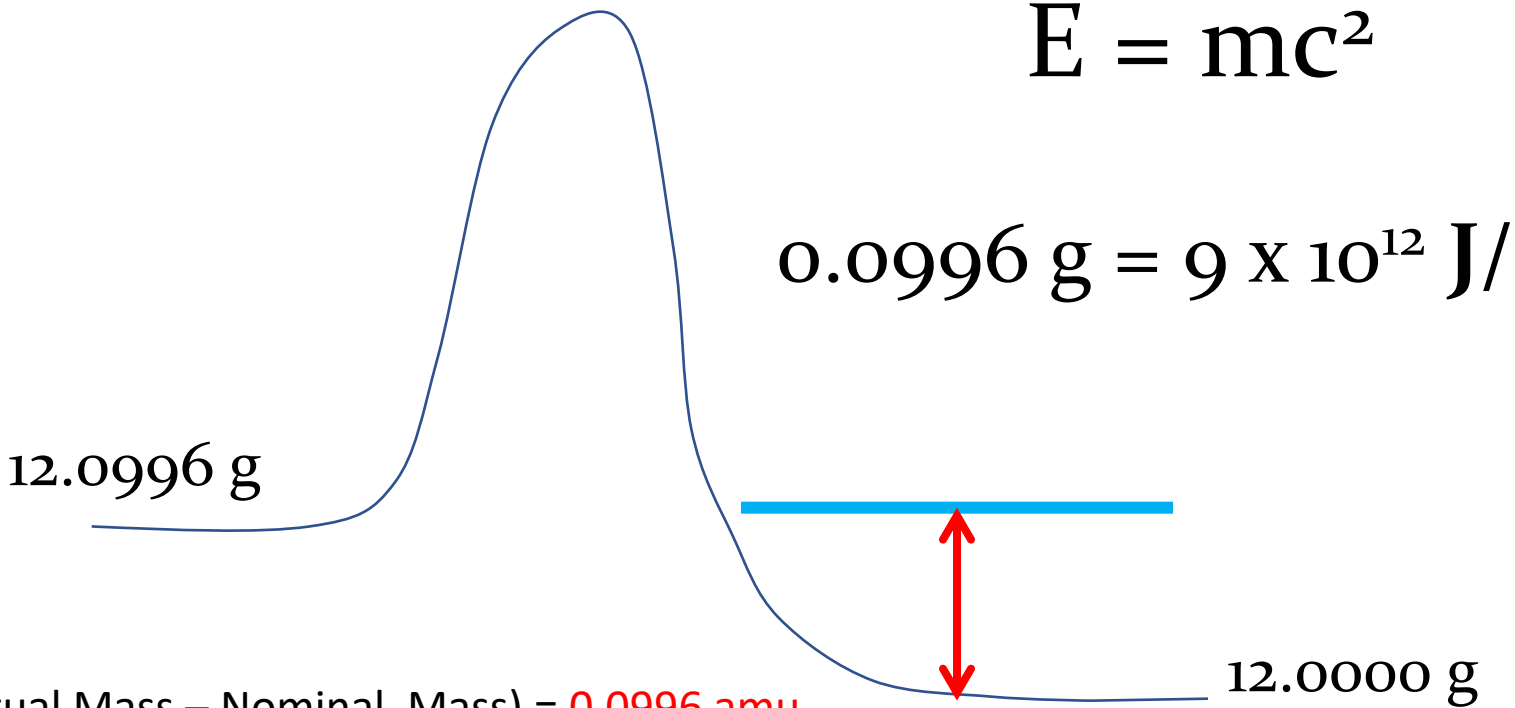
i.e. A carbon atom is lighter than its constituent parts. This Mass Defect is due to the Nuclear Binding Energy.

Nuclear Binding Energy

6p 6n 6e → $^{6}_{12}C$ + Energy (Lots of it!)

$E = mc^2$

$0.0996\ g = 9 \times 10^{12}$ **J/mol**



Mass Defect = (Actual Mass – Nominal Mass) = 0.0996 amu

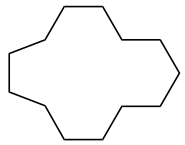
Mass Defect

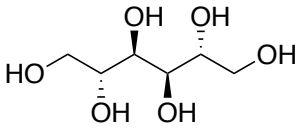
All very interesting, but what does this have to do with mass spec?

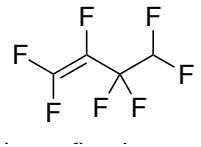
Because the Mass Defect of every isotope is different, accurate mass measurements allow us to identify compounds.

Isotope	Mass (a.m.u.)	Abundance (%)
^{1}H	1.0078	99.985
^{2}H	2.0141	0.015
^{12}C	12.0000	98.89
^{13}C	13.0034	1.11
^{14}N	14.0031	99.64
^{15}N	15.0001	0.36
^{16}O	15.9949	99.76
^{17}O	16.9991	0.04
^{18}O	17.9992	0.20
^{31}P	30.9738	100
^{32}S	31.9721	94.93
^{33}S	32.9715	0.76
^{34}S	33.9679	4.29
^{36}S	35.9671	0.02

Mass Defect







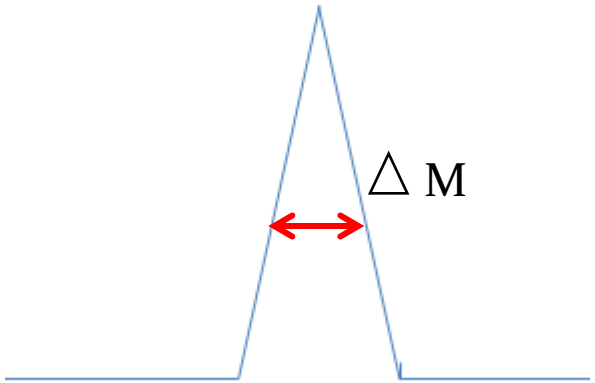
Compound	Formula	Nominal Mass	Accurate Mass
Cyclotridecane	$C_{13}H_{26}$	182	182.2035
Mannitol	$C_6H_{14}O_6$	182	182.0790
Heptafluorobutene	C_4F_7H	182	181.9966

Small differences in mass allow us to tell the difference between the three compounds above. IF we can measure the masses precisely enough.

Resolution

(Mass Spec resolution, different than chromatographic resolution)

- Resolution is a measure of how well the peaks in a mass spectrum are separated from each other



10,000 is regarded as high resolution
1000 is regarded as low resolution

$$\text{Resolution} = \frac{\text{Mass}}{\Delta \text{M}}$$

e.g. if a peak of mass 1000 is 0.1 mass units wide the resolution is :

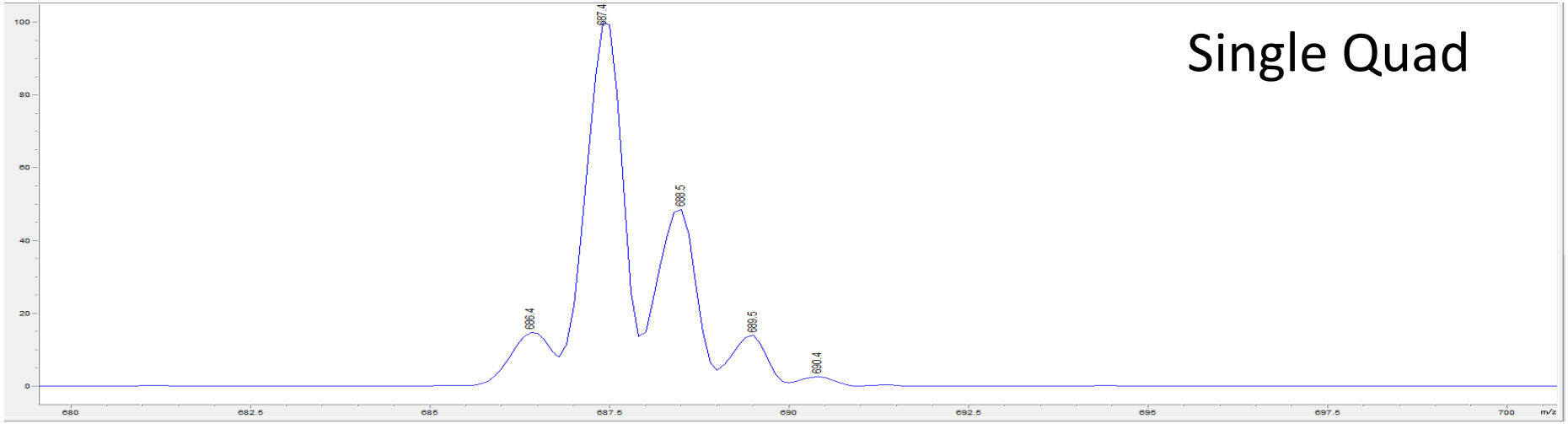
1000/0.1 = **10,000**

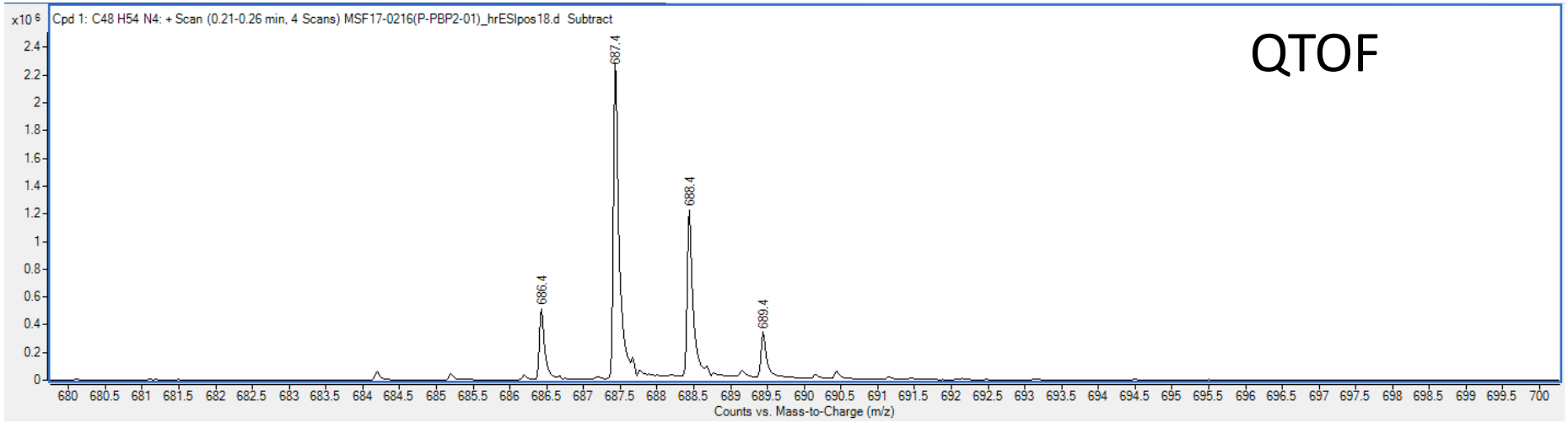
0.1 mass units in 1000 is
$0.1/1000 \times 10^6$ = **100 ppm**

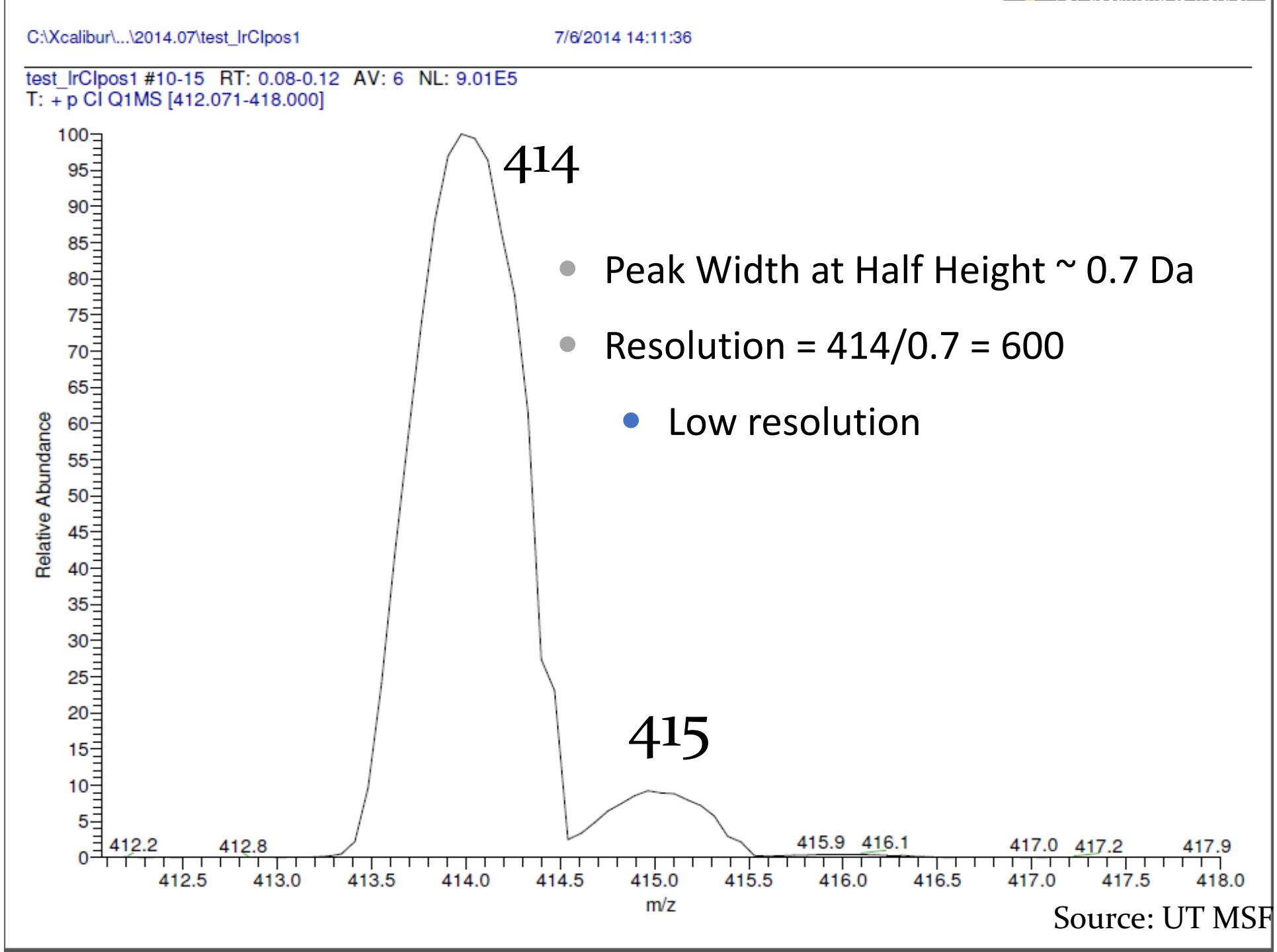
i.e. a 10,000 resolution peak is 100 ppm wide

Low vs. High Resolution Data

+1 charge state

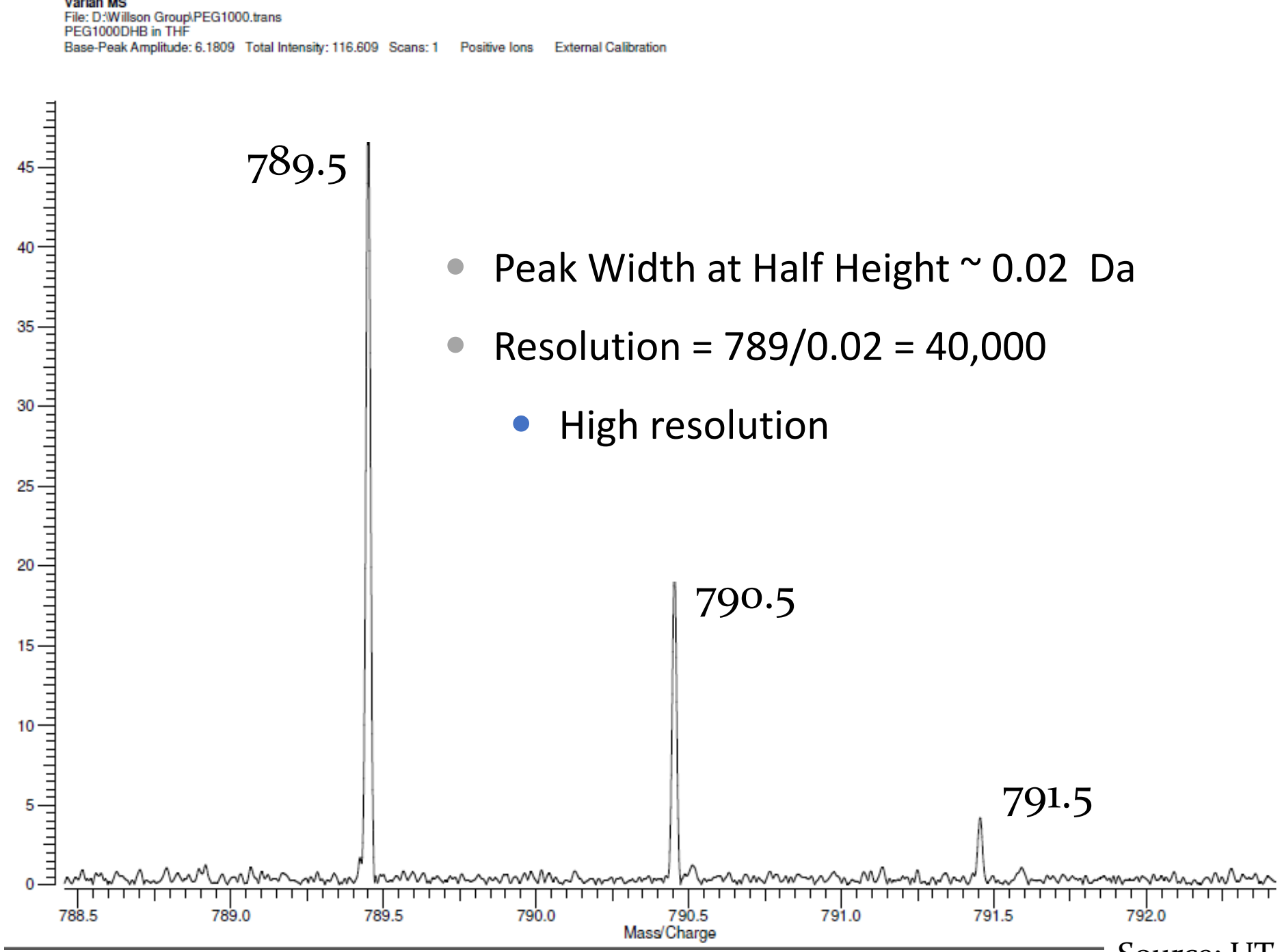






- Peak Width at Half Height ~ 0.7 Da
- Resolution = 414/0.7 = 600
  - Low resolution

Source: UT MSF



- Peak Width at Half Height ~ 0.02 Da
- Resolution = 789/0.02 = 40,000
  - High resolution

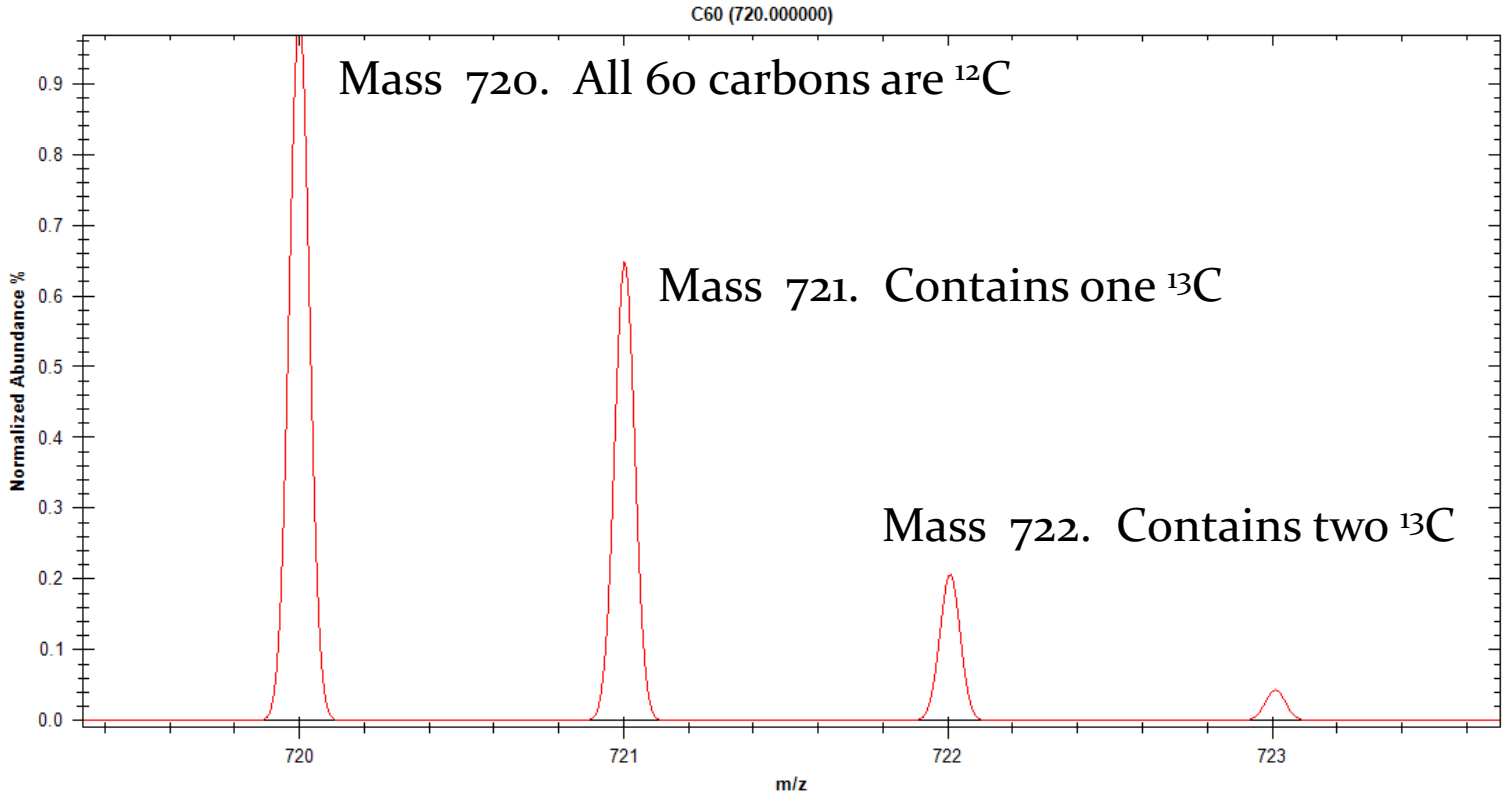
Source: UT MSF

Exact Mass vs. Molecular Mass

- Exact Mass = Monoisotopic Mass is the mass of a molecule that has ONLY the most abundant isotopes in it.
  - i.e. The molecule contains ONLY ^{1}H, ^{12}C, ^{14}N, ^{16}O , ^{35}Cl, ^{79}Br etc.
  - (exact masses can also be calculated for isotopic species variations – but this needs to be specified)

- Molecular Mass = Average Mass is an average of all isotopes.
  - Useful for making up a 1M solution, not useful for mass spec

Mass Spectrum of C_{60}^{+}



- Exact Mass (Neutral Molecule) = 720.0000
  - (Observed C_{60}^{+} is mass 719.99945)
- Average Mass (Neutral Molecule) = 720.64415 Not useful

Mass Error

- An acceptable error for an accurate mass measurement is normally considered to be less than 5 ppm

- A 5 ppm error in measuring the mass of $C_{13}H_{26}$ (mass 182.2035) would be:

$$182.2035 \times 5/1{,}000{,}000 = 0.0009 \text{ Da}$$

So 182.2035 +/- 0.0009 = 182.2026 to 182.2044 is the expected range.

- A result of 182.2100 may seem close but it is not acceptable

- For small molecules the first two decimal places should always be correct

Targeted vs untargeted metabolomics

Untargeted Metabolic Profile

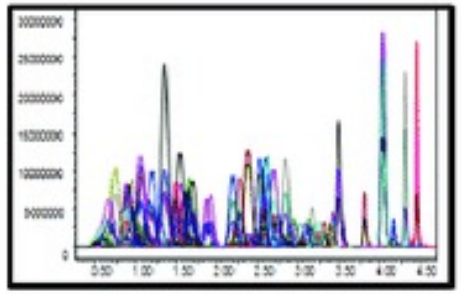
DIMS

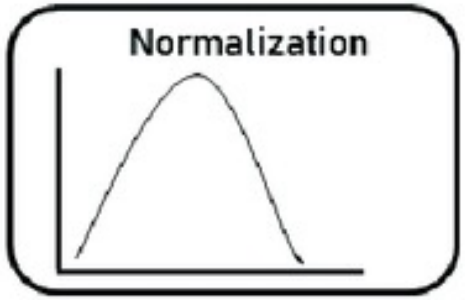
GC-MS | NMR

Analytical tools

LC-MS | DIMS

CE-MS





Putative Metabolites

Pathways Analysis

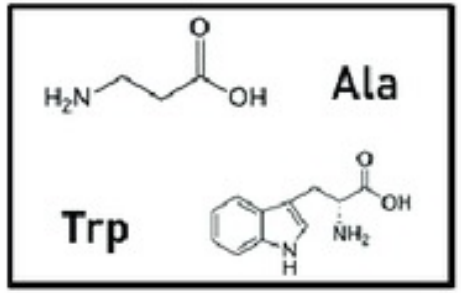
Biological Process

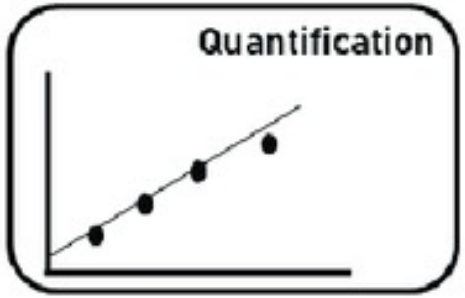
Targeted Metabolic Profile

LC-MS

Analytical tools

GC-MS





Quantification of metabolites

Optimization and Calibration

Validation

Seno de San-martin *Arch Endocrinol Metab* 2021

Untargeted Metabolomics - Data Analysis

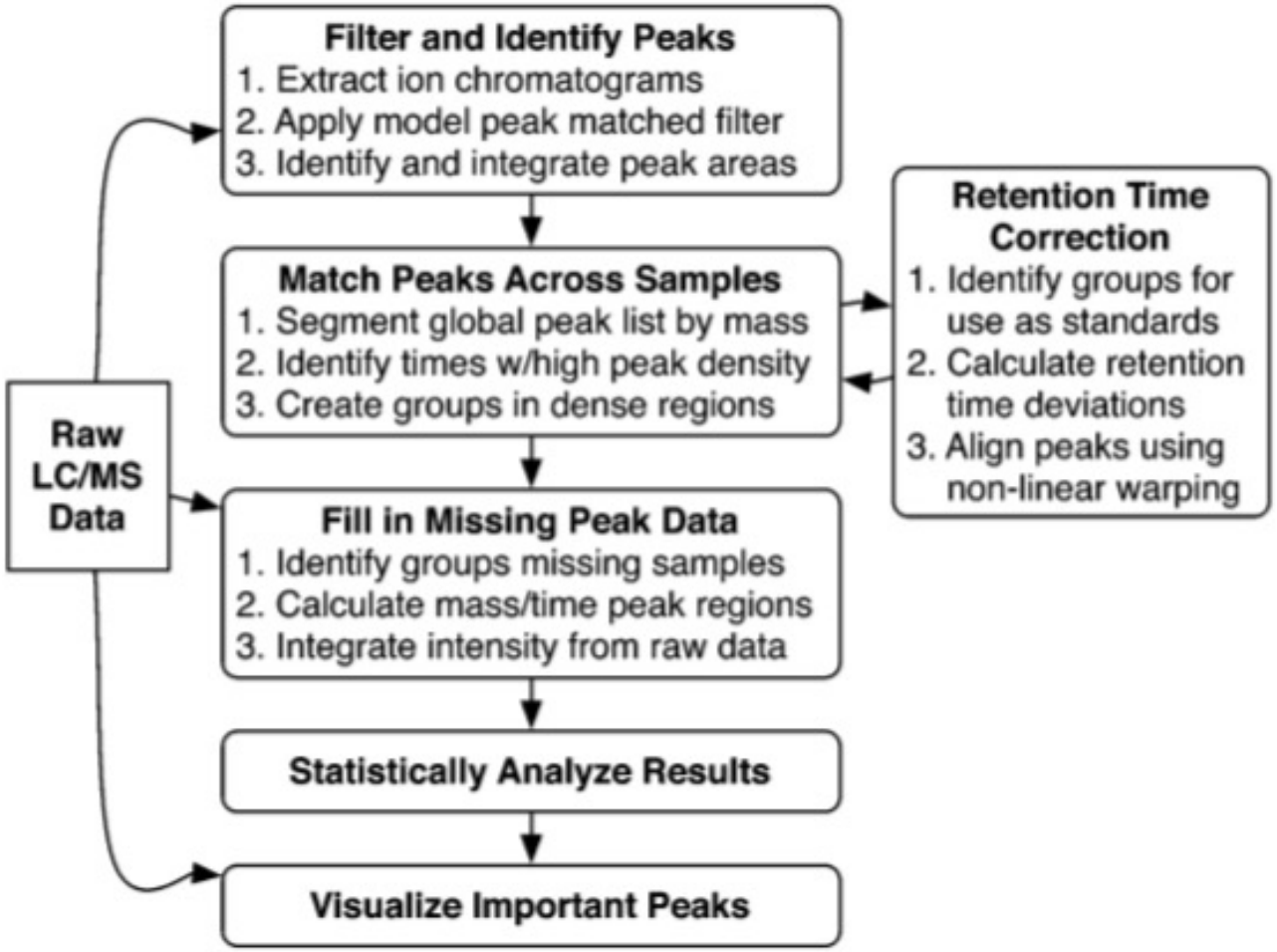
Data preprocessing

Data Annotation

Statistical Analysis

Natural Product Discovery
Biomarker Discovery
Pathway Mapping
Flux

Preprocessing (e.g. Basic XCMS Flowchart)



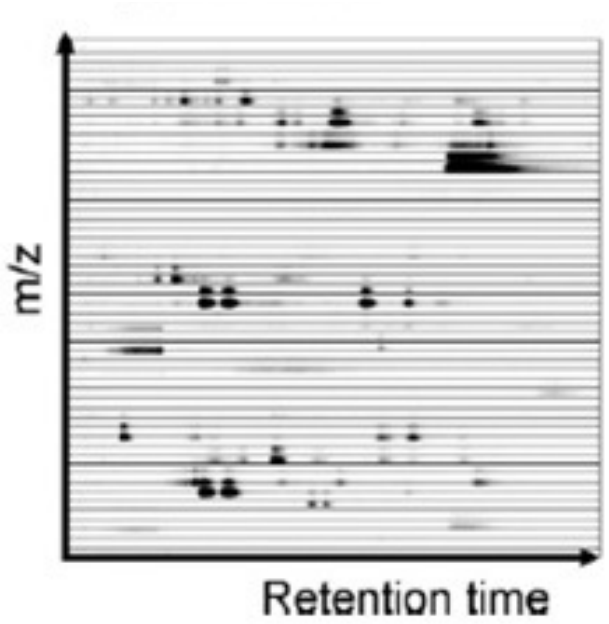
Prepare Input (step 1)

- Various proprietary and open-source formats: e.g. mzXML, mzData, mzML

- MSConvert can be used to convert raw data (http://proteowizard.sourceforge.net/tools.shtml)

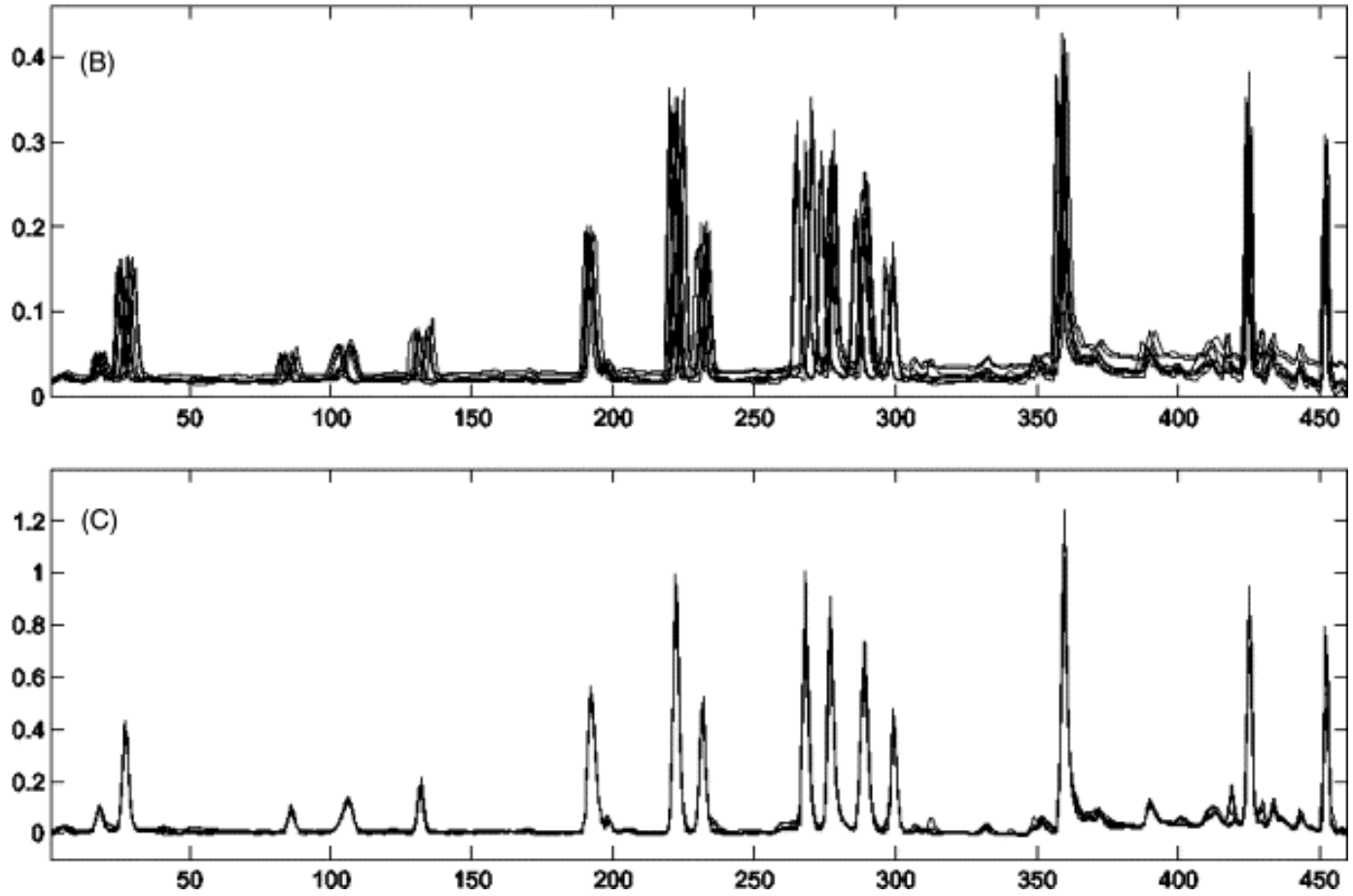
```
# put all .cdf files inside a folder named ‘myspectra’, save the
# folder under your current working directory
> cdffiles <- list.files( ‘./myspectra’, recursive = TRUE, full=T)
# cdffiles now contain absolute path to all raw spectra
> cdffiles
```

Peaking Detection (step 2)



Anal. Chem. **2006,** 78, 779-787

Peak Alignment & Retention Time Correction (step 3)



Filling in Missing Peaks (step 4)

- Missing values are problematic for robust statistical analysis
  1) Peak present in the sample, missed by peak picking
  2) Peak not present
  3) Present, intensity below threshold

- Manual check – use replicate data to fill in peak values

- Re-scan the raw spectra and integrate peaks in the regions of the missing peaks

Results of Peak Detection

- Long list of peaks with
  - mz, mzmin, mzmax
  - rt, rtmin, rtmax
  - peak intensities/areas (raw data)

```
> peaks(xsg)[1:10, ]
            mz mzmin mzmax       rt    rtmin    rtmax       into  maxo sample
 [1,] 200.1000 200.1 200.1 2934.069 2918.299 2948.200  147887.53  9687      1
 [2,] 201.0638 201.0 201.1 2537.824 2522.175 2556.604  204572.42  7726      1
 [3,] 205.0000 205.0 205.0 2789.840 2775.835 2805.384 1778568.94 84280      1
 [4,] 205.9819 205.9 206.0 2791.393 2777.394 2805.384  237993.62 10681      1
 [5,] 207.0821 207.0 207.1 2718.340 2704.553 2732.193  380873.05 18800      1
 [6,] 208.0671 208.0 208.1 2647.286 2631.687 2662.867   96070.72  4112      1
 [7,] 208.1201 208.1 208.2 2716.810 2704.553 2732.193   67967.10  2878      1
 [8,] 219.0848 219.0 219.1 2525.305 2511.220 2540.954  235544.92 11588      1
 [9,] 229.1000 229.1 229.1 2522.175 2509.655 2536.259   87236.08  6649      1
[10,] 233.0390 233.0 233.1 3026.699 3012.052 3041.379  399145.34 19752      1
>
```

Peak Intensity Table

	ko15	ko16	ko18	ko19	ko21	ko22	wt15	wt16	wt18	wt19	wt21	wt22
group	KO	KO	KO	KO	KO	KO	WT	WT	WT	WT	WT	WT
200.1/2926	147887.526	451600.713	65290.3849	56546.1373	85147.0429	162012.439	175177.079	82619.48	51942.9448	69198.2221	153273.469	98144.28
205/2791	1778568.94	1567038.14	1482796.38	1039129.82	1223132.35	1072037.7	1950287.49	1466780.6	1572679.16	1275312.76	1356014.33	1231442.16
206/2791	237993.621	269713.984	201393.416	150107.31	176989.653	156797.035	276541.849	222366.155	211717.713	186850.878	188285.941	172348.755
207.1/2719	380873.049	460629.738	351750.138	219287.968	286848.561	235022.626	417169.585	324892.463	277990.701	220972.352	252874.011	236728.16
219.1/2524	235544.923	173623.378	82364.5911	79480.3984	185792.429	174458.768	244584.471	161184.045	72029.3783	75096.9891	238194.379	173829.951
231/2516	0	70796.2076	222609.068	286232.146	435094.492	100076.188	0	73142.0489	165382.595	240261.212	201316.154	179437.72
233/3023	399145.34	356951.308	410550.655	198416.547	363381.665	317805.788	397107.783	271252.06	334459.927	181901.317	456900.469	294270.588
234/3024	76880.8652	99479.8431	97427.9085	53440.0968	88227.7903	81072.2347	65215.6404	55914.7466	73781.0086	45136.3262	83693.3853	57524.7078
235.1/2694	171995.219	128945.162	155442.477	159288.56	99668.458	112541.211	199981.494	148699.757	156968.295	174559.048	142974.789	106634.911
236.1/2524	252282.035	206031.927	71763.7915	67643.1587	186660.976	198804.282	253791.071	187225.645	53773.0561	90012.6427	256263.261	206486.981
240.2/3681	112440.555	153375.51	193768.652	170641.492	88800.0299	146563.043	122672.854	270872.933	176425.151	187062.973	75235.9352	124070.147
241.1/3679	1465988.67	1318746.57	1215368.74	632037.425	579968.192	561964.932	1468102.64	1594705.06	1006929.31	805533.2	731777.647	522916.137
242.1/3679	280767.61	248792.381	224467.69	109019.428	103855.804	101092.934	280260.239	299453.212	188328.195	139748.3	139968.457	95347.8537
244.1/2832	612169.852	256316.47	90539.8608	35364.5996	54603.1821	115505.42	627834.975	56875.2604	168524.487	27262.6173	110170.73	1352929.86
246.1/2517	27932.12	26061.0198	0	0	56145.3946	47901.5546	105508.752	70508.8223	24386.5748	0	77205.3608	43561.775
249.1/3668	1435000.72	1228148.16	1193346.7	641881.756	536808.256	574005.708	1297985.82	1566269.11	1076654.09	747969.405	688496.327	485427.604
250.1/3668	347794.775	238893.441	248245.61	126623.762	107844.148	118626.972	281745.027	336058.763	215534.93	149872.756	139317.286	99814.4919
254.1/3231	79006.9631	564680.786	89307.0016	127186.81	42745.7444	28747.1784	70817.6569	124200.881	108479.022	42715.7163	78615.5039	44408.536
255.2/3678	1420043.19	1187751.92	1264221.96	673816.219	581177.903	651560.258	1211727.06	1550070.66	1024087.2	875296.292	759075.901	640885.69
256.2/3678	307708.367	263117.303	275104.925	135782.849	121369.162	128540.429	258452.243	338390.94	213618.13	180384.486	164675.68	91822.7078
256.2/3453	229192.802	2899288.97	245869.325	114129.761	65505.2262	62148.0395	249301.23	293543.931	185415.276	98358.9931	170656.944	159465.518

- Peaks are identified by “m/z / retention time”
- Can be directly uploaded to MetaboAnalyst or statistical package of your choice

Untargeted Metabolomics - Data Analysis

Data preprocessing

Data Annotation

Statistical Analysis

Natural Product Discovery
Biomarker Discovery
Pathway Mapping
Flux

Data Annotation

- Resolving LC-MS variables
- Metabolite databases
- Molecular Network Analysis

Example Metabolite Matrix

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	**mzmed**	**248.972489**	**1193.59676**	**341.108204**	**1194.60002**	**1195.60193**	**149.993813**	**377.08437**	**619.295578**	**387.113027**	**533.170449**	**191.054447**	**249.114195**	**149.993554**
2	**MSF19-109_**	16485242	3557089.26	6645550.29	2284059.02	912918.719	4433451.73	968232.066	768839.613	1494469.19	423622.455	378711.944	910674.002	3588272.81
3	**MSF19-110_**	17371407.3	8235996.3	2690115.48	5547637.54	2215462.13	4235876.06	482995.316	1442576.64	632921.003	592973.34	970691.38	961618.125	3862007.33
4	**MSF19-111_**	16110258.2	12013576.5	2124666.29	8020652.88	3310069.87	4431197.77	459836.174	2243191.43	504572.027	248768.652	674150.482	841081.375	3583745.2
5	**MSF19-112_**	17130399.5	7479318.38	5251875.41	5040748.15	2008963.2	3556016.34	649363.453	1492383.46	1194053.55	345130.536	371277.092	988043.06	3877359.76
6	**MSF19-113_**	1391062.31	18971433	5973083.44	12698760.8	5289917.78	550860.916	1258710.71	3210680.72	1364458.01	516965.624	592450.295	2749.04897	4695511.91
7	**MSF19-114_**	18551003.1	8317254	2690734.42	5555245.84	2249767.27	4311580.18	528729.006	1621720.99	621157.475	262563.433	495209.832	1121261.11	3126057.01
8	**MSF19-115_**	15139270.9	44060439.3	4379855.87	29626847.1	12307593.5	4663272.79	347682.572	4970184.93	1045410.61	1028004.47	1945149.55	718667.698	2040166.29
9	**MSF19-116_**	1731679.18	17538661.1	2112794.83	11780391	4922615.22	441545.78	146608.248	2748515.03	507394.408	386604.503	1064070.73	933692.672	4726205.59
10	**MSF19-117_**	18269213.7	4860874.12	3550968.02	3200437.82	1268662.52	4282524.93	1315190.09	1041043.54	797641.901	189138.645	316273.108	1103065.1	3371669.7
11	**MSF19-154_**	19206292.4	3626587.46	3053683.24	2366646.2	939315.882	4397678.65	1078855.53	796526.29	693783.967	123371.579	261564.451	1085614.66	7258748.08
12	**MSF19-155_**	18112310	3328062.88	3483293.44	2158772.14	846785.061	4692732.27	1707277.94	724805.951	782734.461	75951.1873	144918.819	1041448	8676123.94
13	**MSF19-157_**	20022443	4764366.18	2867848.59	3123942.45	1241651.57	4544274.01	1021338.1	1029293.31	663519.845	200031.887	447282.634	1059647.71	3725507.18
14	**MSF19-158_**	20011112	9544684.86	3639557.2	6377602.05	2609968.43	4621445.28	794904.181	1962836.3	870457.348	898748.391	1390047.34	1061547.05	2985635.53
15	**MSF19-159_**	16777230.2	3818171.4	3464166.89	2503336.72	993236.088	4641152.75	768298.622	834331.698	813711.205	243609.294	466213.544	873069.688	3036243.23
16	**MSF19-160_**	20407441.8	4869709.27	1727561.93	3188800.79	1267658.2	4835747.05	296582.961	1007245.84	412070.243	382214.967	1095923.2	1242519.3	3629340.64
17	**MSF19-161_**	18225325.4	5145257.79	1669849.47	3437061.66	1356272.12	4855617.29	300301.306	1082864.72	410803.806	1203795.37	3252872.68	1010455.22	2288684.75
18	**MSF19-162_**	1916598.43	5755546.55	4156049.01	3849136.95	1533431.55	295033.548	541895.581	1147459.33	993542.501	1196488.02	1295068.26	12811.1327	4743083.54
19	**MSF19-182_**	18606381.4	13375326.4	4131577.83	8960557.31	3710051.9	4623354.16	588151.639	2394822.34	1048931.34	583681.791	741679.845	1106676.44	3197993.63
20	**MSF19-185_**	21332746.6	3571433.32	2491658.57	2335242.74	925501.666	4683683.04	246373.768	704736.231	600021.518	81609.206	242967.687	1226118.23	5879367.06
21	**MSF19-186_**	19688503.9	6889502.04	3178510.7	4657464.15	1857568.7	4912090.49	875884.92	1432786.16	767479.043	250394.511	498604.206	1211445.29	7243070.49
22	**MSF19-187_**	20437056.8	14929027.2	3902920.2	9981081.82	4166912.88	4695612.16	231152.638	2629701.12	961666.214	762798.844	1010045.04	1192333.13	3709718.58
23	**MSF19-188_**	1695530.86	5229142.04	3114111.76	3468725.4	1382877.56	540972.03	1773911.19	1011517.94	718114.375	128686.432	203055.607	1262.01344	4660607.14
24	**MSF19-189_**	1955878.82	34257124.5	3381336.6	23016258.7	9518257.56	532472.603	291305.581	4058289.72	831541.126	867606.669	1205815.39	1212084.66	4792051.12
25	**MSF19-190_**	21219305.3	4679190.8	5194788.91	3039028.3	1223286.62	4543007.22	1413539.53	894395.89	1203891.32	824008.701	925710.15	1194585.42	3901885.07
26	**MSF19-191_**	19661726.9	5705846.96	3037844.09	3833894.37	1484807.51	4687634.98	533273.608	1004638.25	705610.307	153600.595	376994.4	1122170.25	3847543.12
27	**MSF19-192_**	20230533.3	3273956.42	3484267.7	2131909.65	858444.605	4644967.12	475588.149	719847.138	836679.856	234282.075	404890.487	1181827.9	3710548.28
28	**MSF19-193_**	21995555.9	4868125.6	3858390.37	3187050.33	1269270.93	4493606.78	1343311.93	1030870.57	904174.841	538323.674	851741.267	1265569.66	9459197.48
29	**MSF19-194_**	13657653.4	4574431.75	5428734.67	3002875.11	1191909.86	4933176.7	1270426.07	684971.998	1038926.03	685422.251	679016.76	612408.578	3545969.06
30	**MSF19-195_**	17189985.1	943063.391	3408219.62	624296.116	249505.045	3542134.15	772994.203	243196.501	787731.22	44809.7259	125450.795	838915.984	4026150.53
31	**MSF19-196_**	15146333.5	4195376.4	2924492.85	2748508.38	1084831.1	4835703.15	268210.651	861033.75	707323.203	221516.485	528880.124	811834.346	3434970.79
32	**MSF19-197_**	17629550.8	8117408.86	3871939.11	5414117.05	2198926.49	4517724.6	268606.518	1606055.82	934002.827	566598.089	812654.006	1019065.47	6085569.53
33	**MSF19-198_**	18152147.6	8681688.1	2412609.32	5762846.46	2344898.21	3429791.41	200568.854	1460787.25	602188.997	376586.583	810631.346	1038041.11	3940420.95

Example Annotated Matrix

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	featureidx	MSMS	mzmed	mzmin	mzmax	rtmed	rtmin	rtmax	maxint	isotopes	adducts	peakgroup	METLIN	name	METLIN_MSMS		1827	unknowns		
2	640		395.026855	395.021621	395.030829	0.37516667	0.33833333	0.79533333	14495.6436			4	0::::M-H::1046678::NA::|		y		0.63569937	% unknown (no MTELIN hit)		
3	600		482.966325	482.963611	482.978975	21.47875	21.1543333	21.648	11449.335			18	0::::M+Cl::1050341::NA::|		y					
4	2260		527.174422	527.168254	527.176391	2.18483333	2.06066667	2.3265	2208.95068			166	0::Asp Glu Glu His::M-H::		y					
5	414		310.999912	310.992369	311.002392	21.5303333	21.2766667	21.9353333	16521.1191			2	0::Dihydroferulic acid 4-su		y					
6	2494		759.169428	759.162092	759.174562	13.7743333	13.2193333	14.0206667	2070.70557			22	0::Pelargonidin 3-(4'''-p-c		y					
7	718		399.057142	399.051007	399.062668	0.35916667	0.34	0.42983333	12955.5244			4	1::::M-H:: 996739::NA::| 1		y					
8	2211		479.245171	479.240391	479.2488	3.69625	3.666	3.76083333	5067.02637			7	1::::M-H::1018963::NA::|		y					
9	1864		301.090957	301.086276	301.093257	0.541	0.50466667	0.62966667	7151.35547			74	1::::M-H::1032495::NA::|		y					
10	2091		341.133314	341.120657	341.137681	4.091	4.0495	4.374	4745.13379			93	1::::M-H::1039534::NA::|		y					
11	453		443.055497	443.053163	443.05872	2.47733333	2.3735	2.774	55555.8633	[95][M]-		3	1::::M-H::1050175::NA::|		y					
12	1218		507.071697	507.067943	507.074636	3.2525	3.19566667	3.277	17133.8262			15	1::::M+Cl:: 992849::NA::|		y					
13	1334		497.060926	497.055604	497.064842	3.30775	3.262	3.34416667	22031.3848			14	1::::M+Cl::1001655::NA::|		y					
14	1627		499.083815	499.078268	499.096465	4.43466667	4.3935	4.45366667	14953.1738			48	1::::M+Cl::1007617::NA::|		y					
15	2144		409.992637	409.990863	409.994606	15.0896667	14.7695	15.5421667	2585.9043			16	1::::M+Cl::1084185::NA::|		y					
16	1758		637.205143	637.188447	637.212947	0.4125	0.37316667	0.50716667	5697.1167			4	1::5,4'-Dihydroxy-6-C-prer		y					
17	56		901.478302	901.473694	901.481404	4.97191667	4.9385	4.991	228211.375	[309][M]-		23	1::PG(22:6(4Z,7Z,10Z,13Z		y					
18	2130		391.206471	391.204141	391.207938	12.13	11.9731667	12.3866667	3393.73364			140	2::::M-H:: 992845::NA::| 2		y					
19	1442		415.034603	415.025612	415.047831	0.35966667	0.34116667	0.46983333	5852.31543			4	2::::M-H::1050731::NA::|		y					
20	437		431.002411	430.988508	431.006951	0.35983333	0.34	0.443	27520.9356			4	2::::M-H::1074166::NA::|		y					
21	2093		343.10208	343.100594	343.104854	2.32383333	2.23566667	2.38716667	3007.60352			162	2::::M-H::1085560::NA::|		y					
22	363		477.076326	477.063329	477.087304	0.37666667	0.32116667	0.78583333	26293.5352	[111][M]-		4	2::::M+Cl::1056746::NA::|		y					
23	1440		471.024179	471.020415	471.02647	3.746	3.70933333	3.77666667	12510.7217			12	2::::M+Cl::1066819::NA::|		y					
24	885		433.00022	432.991433	433.003587	0.3595	0.34083333	0.443	11328.7461			4	2::::M+Cl::1083169::NA::|		y					
25	416		164.92698	164.925907	164.930461	21.782	21.498	21.9425	13610.2334			37	2::Calcium formate::M+Cl		y					
26	1215		374.001083	373.990785	374.005608	21.454	21.1488333	21.6443333	5786.1875			18	2::FLUQUINCONAZOLE::M		y					
27	1789		279.162493	279.15936	279.163856	9.85141667	9.593	10.056	2679.9043			165	3::::M-H::1023223::NA::|		y					
28	755		493.98385	493.980535	493.98683	21.4535	21.1488333	21.632	8905.54981			18	3::::M+Cl:: 995469::NA::|		y					
29	2188		451.044447	451.03979	451.047943	0.73033333	0.68716667	0.82733333	7071.48438			60	3::::M+Cl::1033575::NA::|		y					
30	1203		635.191784	635.18849	635.195327	3.7285	3.51433333	3.9275	9884.02734			7	3::100-2::M+Cl::63806::NA		y					
31	1000		667.254359	667.24456	667.259997	4.0295	3.98866667	4.055	27078.4414			28	3::Asp Trp Trp Tyr::M-H::1		y					
32	761		245.039574	245.028996	245.041979	0.37766667	0.34333333	0.46516667	14235.5811			4	3::Biphenylene sulfone, 3		y					
33	2248		513.30593	513.301859	513.308716	14.8996667	14.6156667	15.178	1632.40625			16	3::Lys Leu Pro Ala Ser::M-		y					

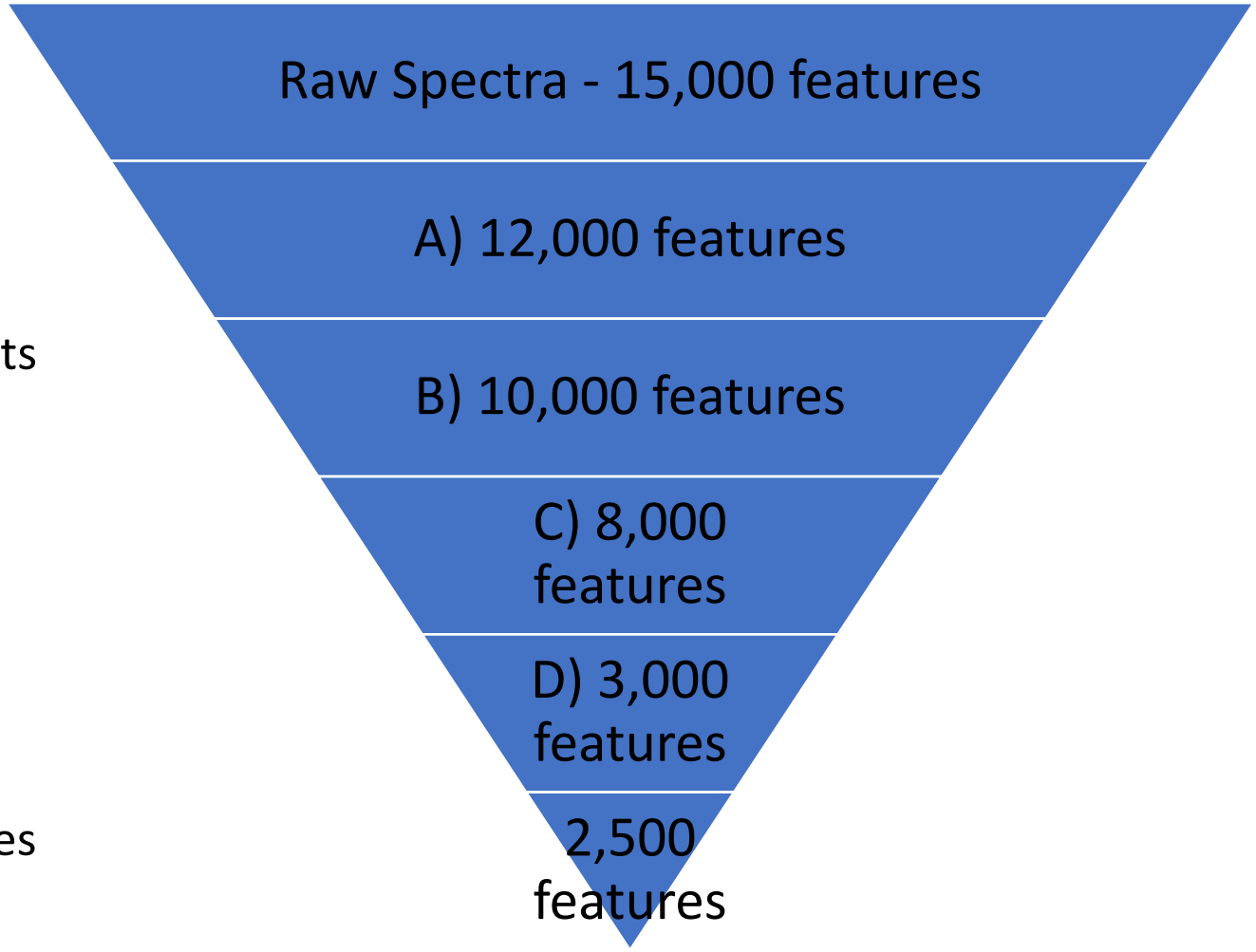
Resolving LC-MS variables

A) Identify, remove (or consolidate) adducts and remove multiple charged species

B) Identify, remove (or consolidate) fragments (neutral losses, breakdown products, rearrangements)

C) Identify, remove (or consolidate) isotope peak

D) Remove noise peaks (from sample blanks or peaks that do not appear in >2/3 technical replicates or peaks that do not show dilution trends in 4 dilution replicates



Metabolite Identification

Identification confidence	*Minimum data requirements*
1) Positively identified – confirmed by match to known standard	MS; MS2; RT; Reference standard
2) Putatively identified – match to MS, MS/MS or RT library	MS, MS2; Library MS, MS2; MS; MS2 Exp. data
3) Putatively identified to compound class	MS; MS2; Exp. Data
4) Unequivocal molecular formula	MS isotope/adduct
5) Exact mass of interest	MS

Schymanski, E. *et al. Environ. Sci. Technol.* **2014**, *48,* 2097
Summer, I.W. *et al*. *Metabolomics* **2007**. *3*, 211

Metabolite Databases

Points of Caution

- Many databases do not distinguish source in library (synthetic, plant, human, drug, etc. all mixed)
- Matches do not make sense
- If you know sample source – select appropriate database (HMBD, FooDB, DrugBank, KnapSack, etc.)

Metabolite Databases

- T3DB (~4000)
- DrugBank (~1300)
- FooDB (~30,000)
- DrugBank (~1600)
- HMDB (~20,000)
- METLIN (~500,000)
- SuperNat
- PubChem
- ChemSpider

- Toxins/Environmental Chemicals
- Drug metabolites
- Food additives/phytochemicals
- Drugs
- Endogenous metabolites
- Reference Library
- Natural product focus
- Includes synthetic molecules
- Includes synthetic molecules

Molecular networking (GNPS)

- Link features to each other into a network based on MS2 spectra and *m/z* differences
- Can obtain info about a feature's chemical class via guilt-by-association (MolNetEnhancer)

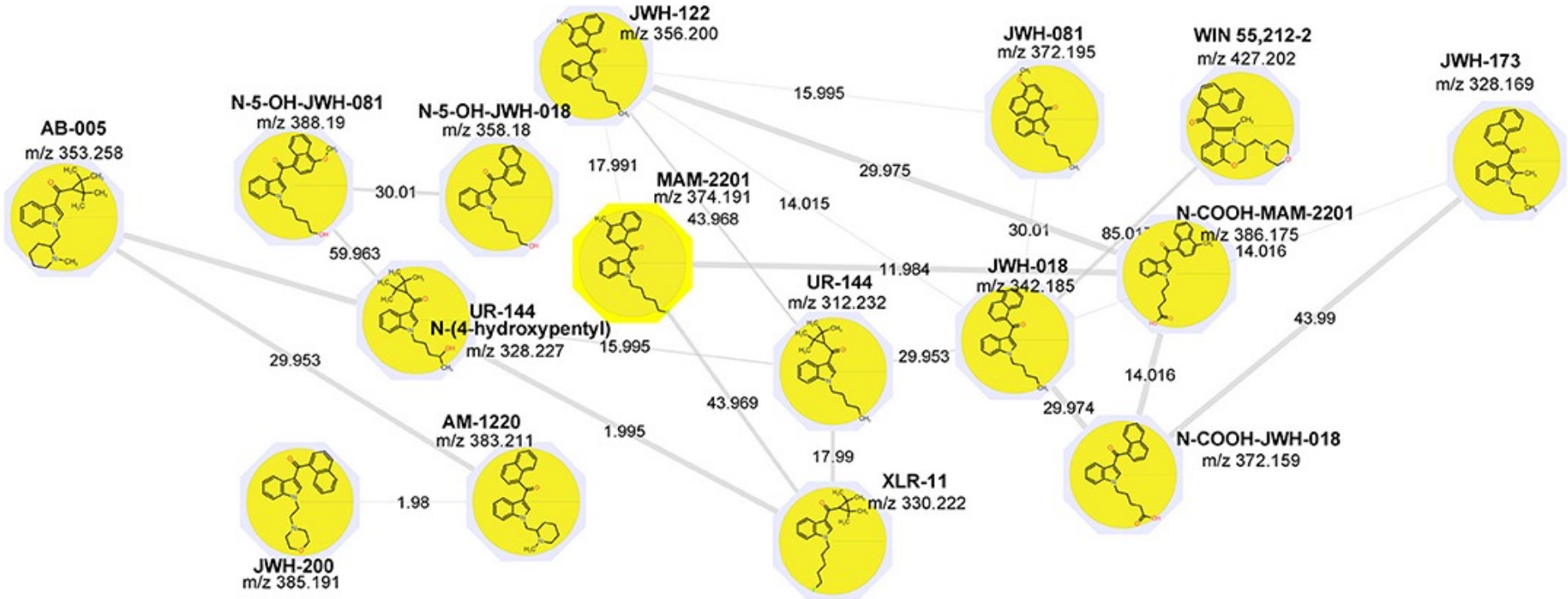
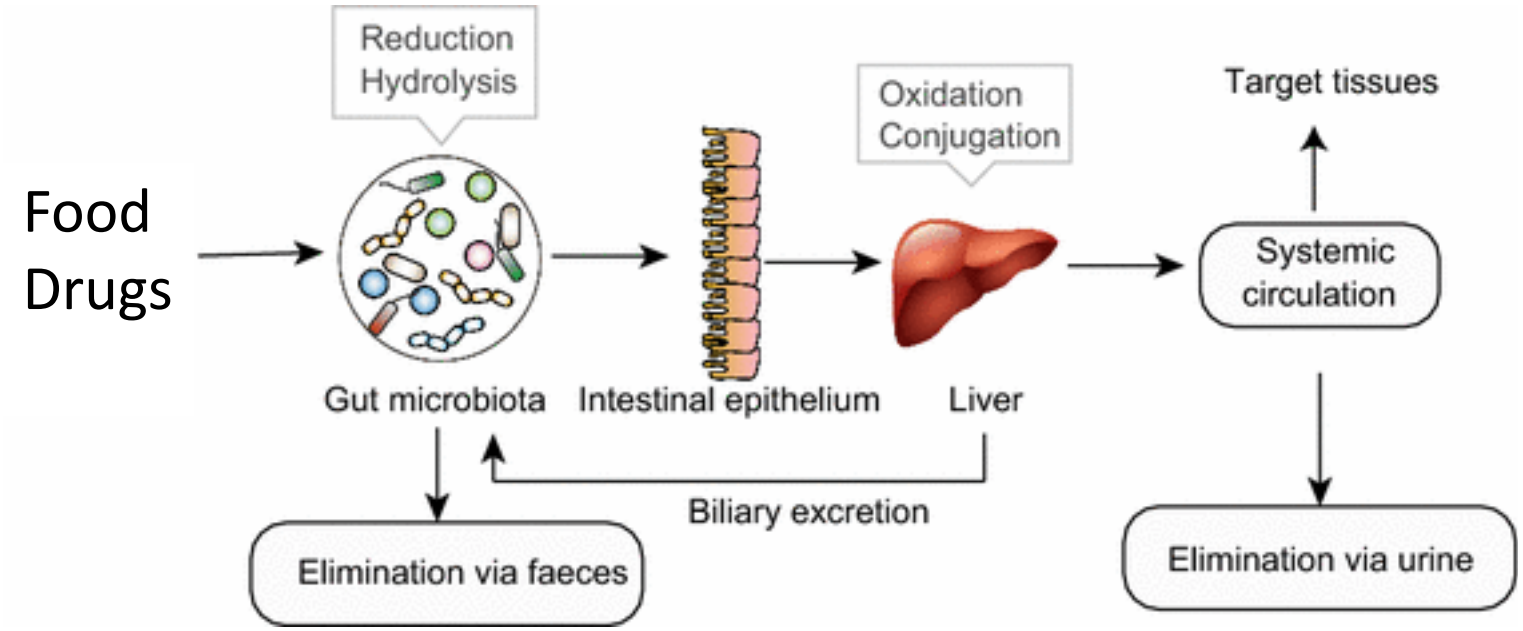


Image from Vincenti *Frontiers in Chemistry* 2020

Today's data

- 40 urine samples from the MicroPD study
- Untargeted metabolomics (TripleTOF)



Doestzada Protein & Cell 2018