

Mass spectroscopy for proteins

Dr Manaf Abdulrahman Guma
University Of Anbar- College Of Applied Sciences-Hit(Heet)
Department Of Applied Chemistry

1

What does mass spectroscopy check?

- Mass spectroscopy studies the proteomics.
- What are the Applications of the proteomics?
 1. Protein identification – Specifically complex protein mixtures.
 2. Study PTM post translational modification of the protein such as phosphorylation, acetylation a methylation etc.
 3. protein molecular weight determination.
 4. Structural studies by H/D exchange.

2

What does proteomics study?

- introduction to proteomics
- protein separation techniques
- proteolytic digestion
- peptide mass fingerprinting
- LC-MS/MS

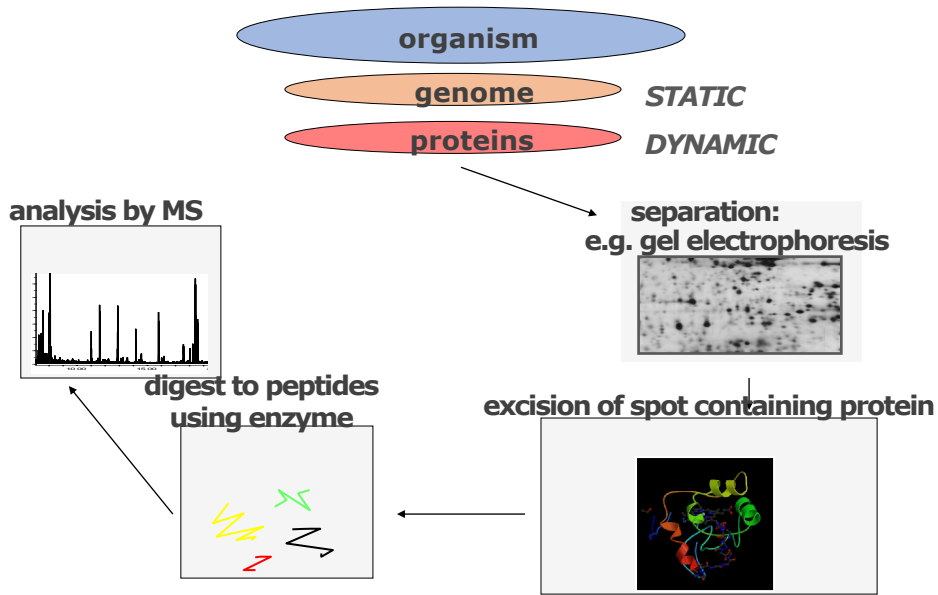
3

What is proteomics?

- The Identification AND Quantification of the full protein complement expressed by the genome of an organism at a particular point in time.
- It studies:
 1. It finds which protein is present in the sample.
 2. How much proteins are present in the sample.
 3. Which gene the protein is belong to?

4

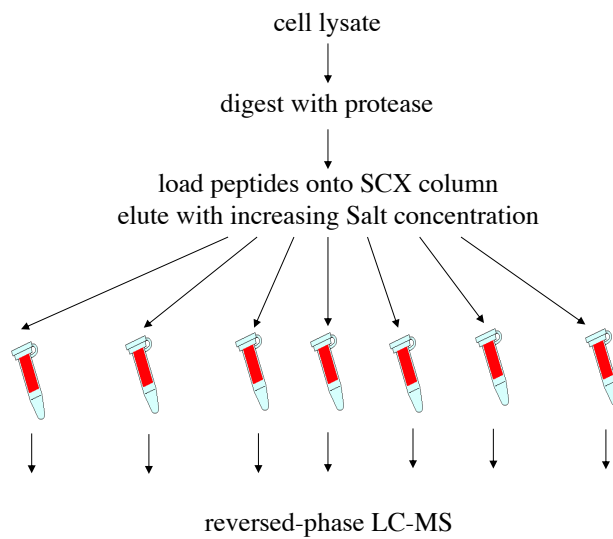
What is the strategy for proteomic analyses?



5

separation techniques

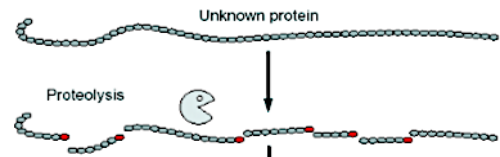
multidimensional HPLC for separating Large numbers of proteins



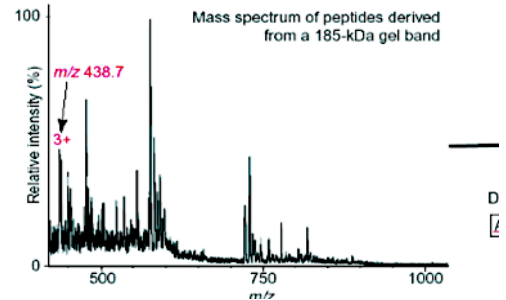
6

What is proteolytic digestion?

- It is a method to digest the protein by enzymes or chemical to make small fragments.



Record mass spectrum of the peptide mixture

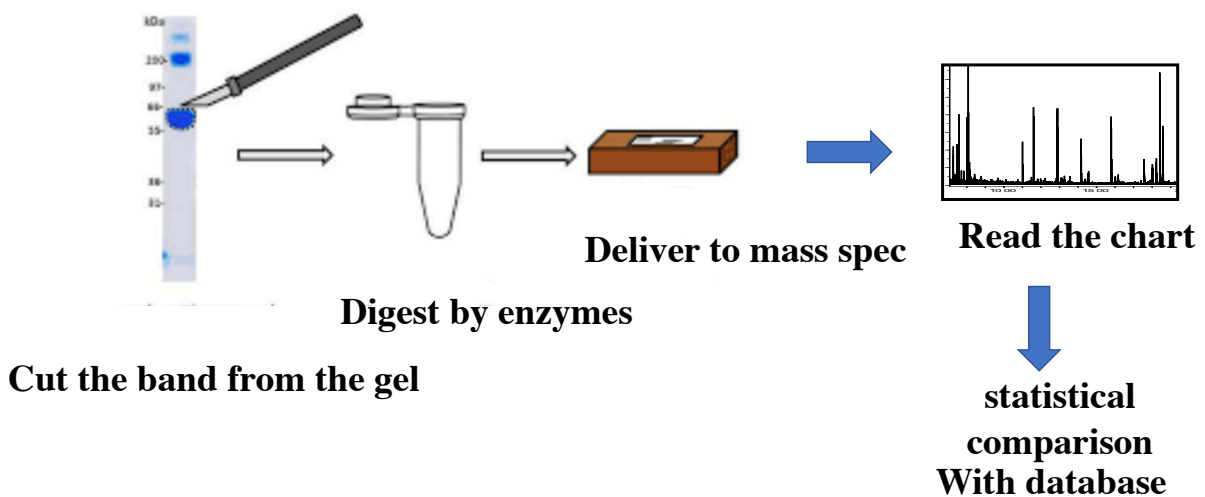


Mass spectrometric identification of an unknown protein can be obtained from determining the sequence of proteolytic fragments and comparing the sequence with available databases

	cleavage sites	exceptions
<u>Trypsin</u>	C-term of K or R	if P is C-term of K or R
<u>Chymotrypsin</u>	C-term of F , Y , W , (M), (L)	if P is C-term of F , Y , W
<u>CNBr</u>	C-term of M	
Lys-C	C-term of K	
Arg-C	C-term of R	if P is C-term of R
Asp-N	N-term of D	
Glu-C (bicarbonate)	C-term of E	if P or E is C-term of E
Glu-C (phosphate)	C-term of D or E	
Pepsin (pH 1.3)	C-term of F , L	
Pepsin (pH >2)	C-term of F , L , W , Y , A , E , Q	
Proteinase K	C-term of F , L , W , Y , A , V , I	

7

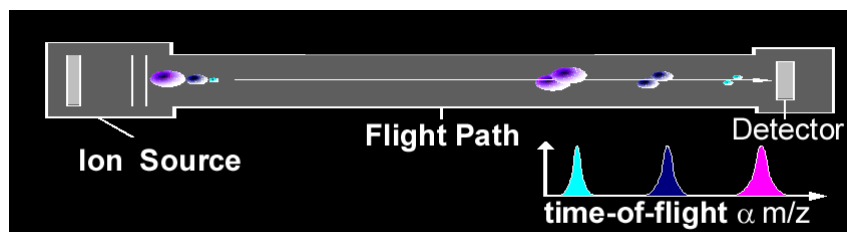
What are the proteolytic digestion steps?



8

How does mass work?

Mass determination by Time-of-Flight

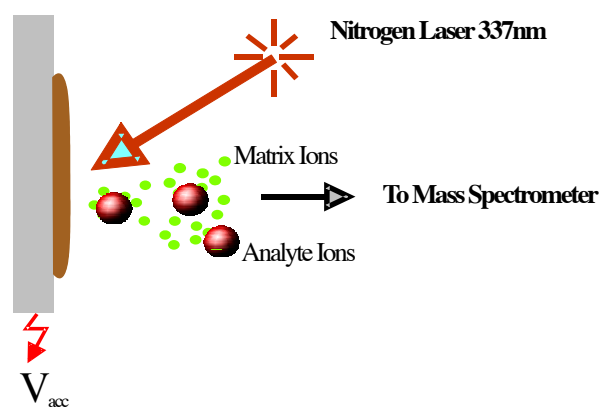


1. ions generated in source
2. accelerated using high voltage
3. drift to detector
4. time-of-flight inversely proportional to m/z

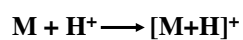
9

What is used to ionize the molecules?

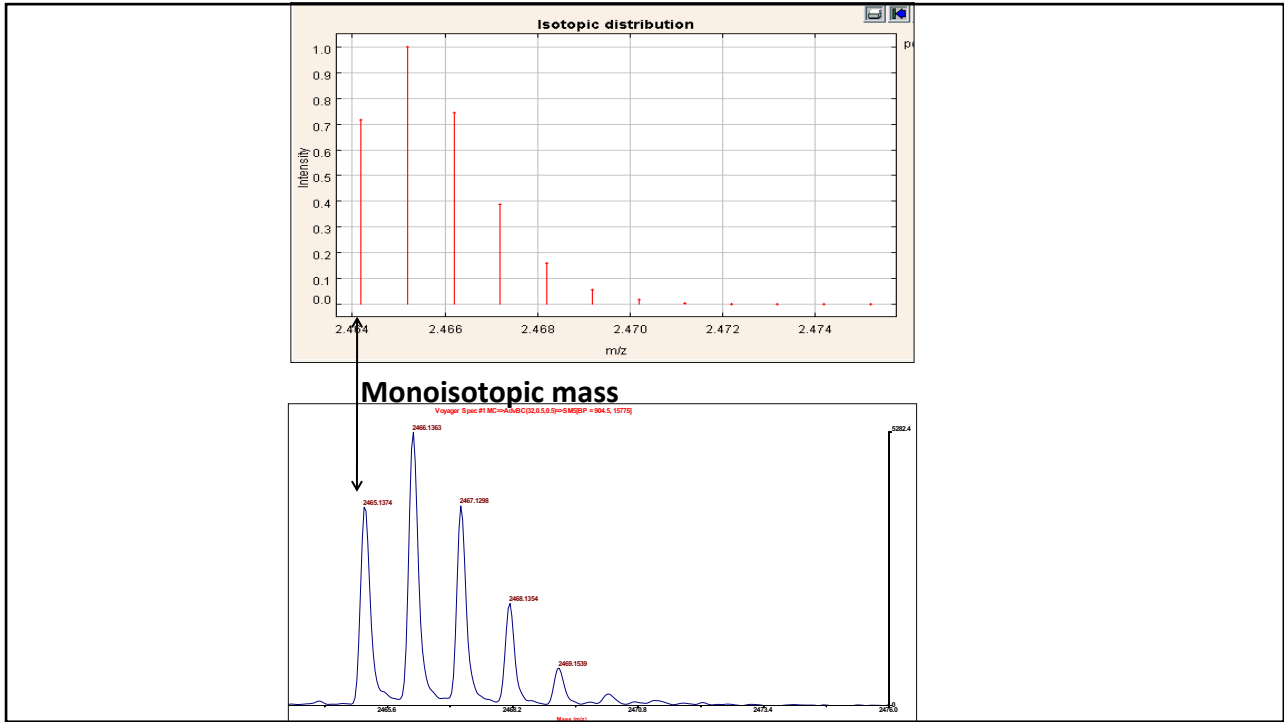
Laser Desorption/ionisation



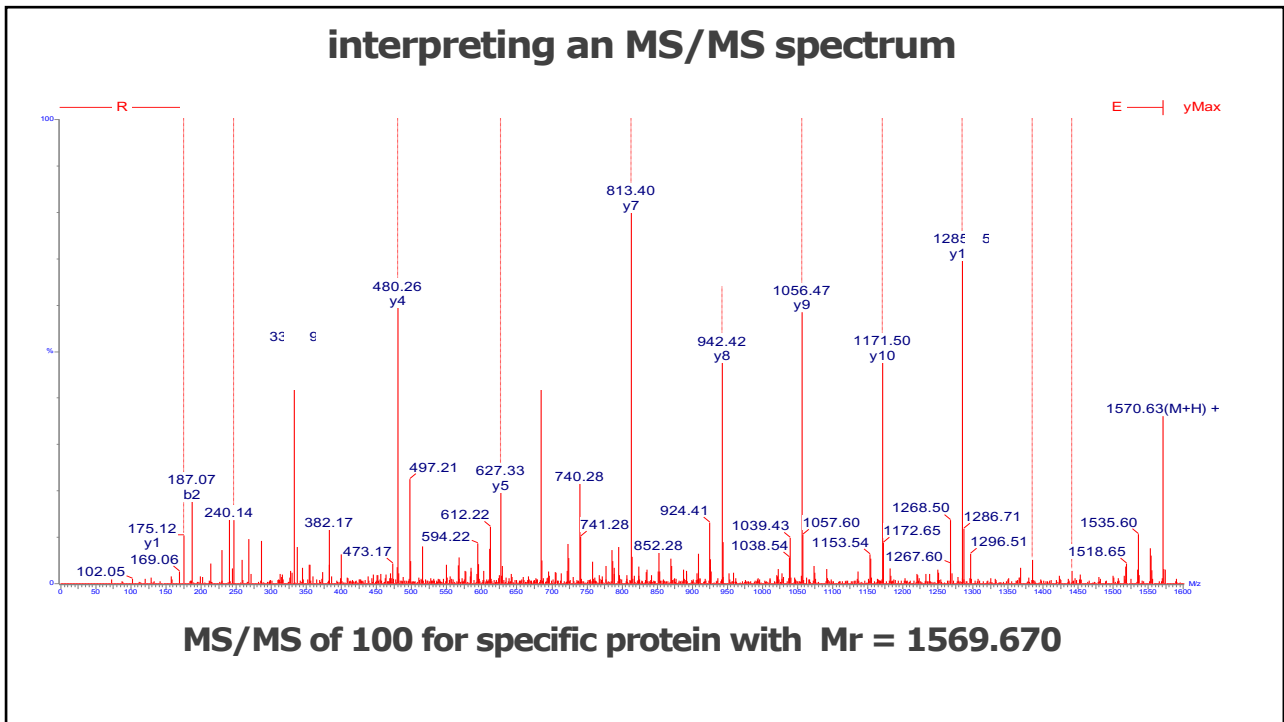
Charge transfer occurs between Matrix Ions and Analyte molecules



10



11



12

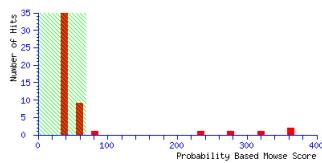
How do we do statistical comparison With database? By searching of PMF data which are online available.

(MATRIX) Mascot Search Results

User : Andrew Bottrill
 Email : arb29@le.ac.uk
 Search title : BSA_03_0001.dat - Sample Info, D:\Andrew\Data\Voyn
 Database : NCBInr_20060211 (3292813 sequences; 1128164434 re
 Taxonomy : Mammalia (mammals) (445638 sequences)
 Timestamp : 12 Feb 2006 at 13:38:40 GMT
 Top Score : 363 for **gi|162648**, albumin [Bos taurus]

Probability Based Mowse Score

Protein score is $-10 * \log(P)$, where P is the probability that the observed match is a random match.
 Protein scores greater than 69 are significant ($p < 0.05$).



Observed	Nr(expt)	Nr(calc)	Delta	Start
712.4040	711.3967	711.3664	0.0304	23
847.5360	846.5287	846.4963	0.0324	24
898.5210	897.5137	897.4742	0.0395	483
927.5470	926.5397	926.4861	0.0536	164
1068.4870	1067.4797	1067.4342	0.0455	413
1138.5540	1137.5467	1137.4907	0.0561	499
1163.6820	1162.6747	1162.6233	0.0514	64
1166.5490	1165.5417	1165.4856	0.0561	466
1193.6580	1192.6507	1192.5949	0.0559	23
1249.6760	1248.6687	1248.6138	0.0549	33
1283.7790	1282.7717	1282.7033	0.0684	363
1305.7730	1304.7657	1304.7089	0.0569	403
1339.7830	1338.7757	1338.6853	0.0495	263
1419.7900	1418.7827	1418.6864	0.0564	89
1439.8760	1438.8687	1438.8044	0.0643	366
1443.7030	1442.6957	1442.6347	0.0610	284
1461.6480	1460.6407	1460.5817	0.0591	74
1479.6570	1478.6497	1478.5961	0.0616	423
1502.6860	1501.6787	1501.6065	0.0723	378
1532.6930	1531.6817	1531.6138	0.0579	298
1554.7280	1553.7207	1553.6456	0.0751	387
1567.6110	1566.6037	1566.5254	0.0693	347
1638.5940	1637.5867	1637.5004	0.0593	437
1747.7640	1746.7567	1746.6977	0.0590	184
1880.5840	1879.5767	1879.5138	0.0629	508
1907.5930	1906.5857	1906.5135	0.0702	523
1927.6740	1926.6667	1926.5910	0.0718	583
2020.0370	2019.0297	2018.9619	0.0679	133
2248.0120	2247.0047	2246.9354	0.0693	247

No match to: 856.5720, 861.1200, 1249.8670,

(MATRIX) Mascot Search Results

Protein View

Match to: **gi|162648** Score: 363 Expect: 2.2e-31
albumin [Bos taurus]

Nominal mass (M₀): 71244; Calculated pI value: 5.82
 NCBI BLAST search of **gi|162648** against nr
 Unformatted [sequence string](#) for pasting into other applications

Taxonomy: [Bos taurus](#)
 Links to retrieve other entries containing this sequence from NCBI
[gi|1351907](#) from [Bos taurus](#)

Fixed modifications: Carbamidomethyl (C)
 Variable modifications: Oxidation (M)
 Cleavage by Trypsin: cuts C-term side of KR unless next residue is
 Number of mass values searched: 36
 Number of mass values matched: 29
 Sequence Coverage: 53%

Matched peptides shown in Bold Red

```

1 MKVVTFTISLL LFFSSAYSRG VFRFDTHKSE IAHREKDLGE EHWGLVLIA
51 FSQYLQCCPF DEHVKLVEL TEFAKTCVAD ESHAGCEKSL HTLFGDELCK
101 VASLRETYGD MADCCERQEP ERNECFLSHK DSDPDKPKL PDPTLCKEP
151 KADKKKFWGK YLVEIARRHP YFYAPLLYY AMKNGVQVE CCQAEKDGAC
201 LLPKIEITMRE KVLASSARQR LRCASIQFG ERALKAWSYA RLSQKPKKAE
251 FVEVTKLWTD LTKVHKECCH GLLECADRR ADLAKYICDN QDTSISKLEK
301 CDDKPLLEKS HCIAEVEKDA IPENLPLTA DFAEKDVCK NYEAKDAFL
351 GSPLYEYSRR HPEAVSVLL RLAEVPEALT ECCAKDDPH ACYSTVDFDKL
401 KULVDEPQHL IKQNCDFEKL LGEYGFQNAL IVRYTRKVPQ VSTPTLVEVS
451 RSLGRVGTRC CTKPESEKRP CTEDYLSLIL NRLCVLHEKT PVSEKVTKCC
501 TESLVNRRPC FSALTPDETY VPKAFDEKLF TFHADICTLP DTERQIKKQT
551 ALVELLEKRP KATEEQLKIV MENFVAFVCK CCAADREKAC FAVEGPKLVV
601 STQTALA
    
```