

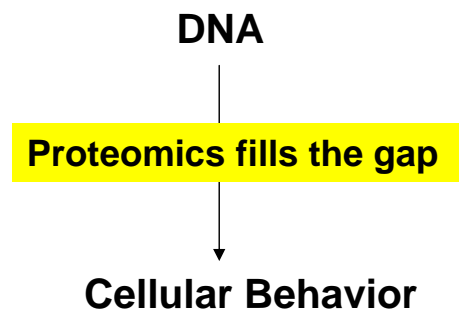
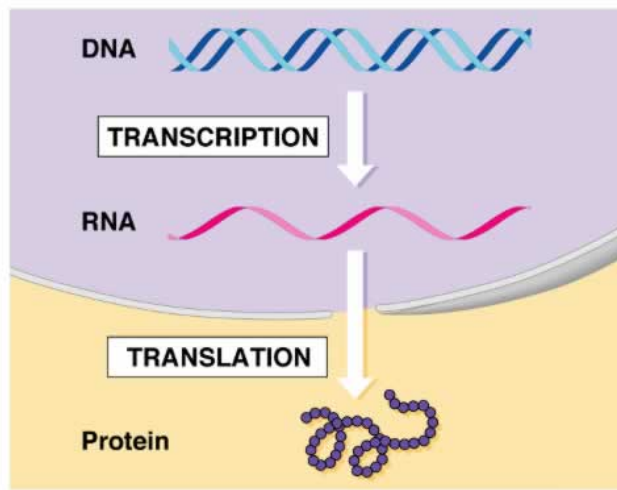
Introduction to Mass Spectrometry Based Proteomics

Mass Solutions Technology

Outline

- **Mass Spectrometry**
 - Ion source
 - Mass Analyzer
 - Tandem mass spectrometry
- **Protein Identification**
 - Peptide mass mapping
 - LC-MS/MS for protein ID

Proteomics

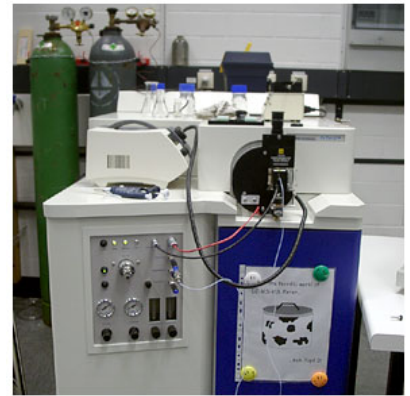
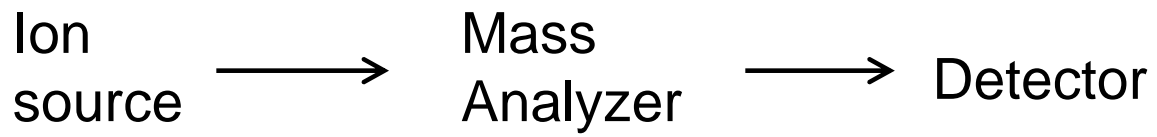


Functional protein

Mass Spectrometry

1. Ion Source
2. Mass Spectrometry
3. Tandem mass spectrometry

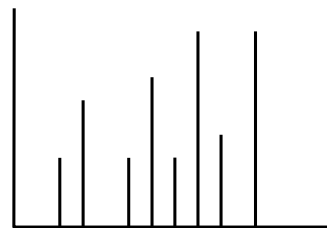
The composition of mass spectrometry



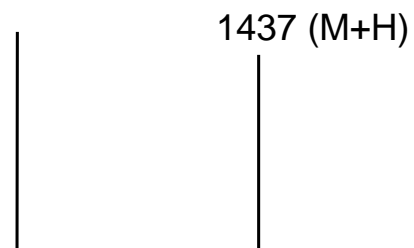
Ion source

Peptide:
"FESNFNTHATNR"
MW: 1436

Hard ion source →

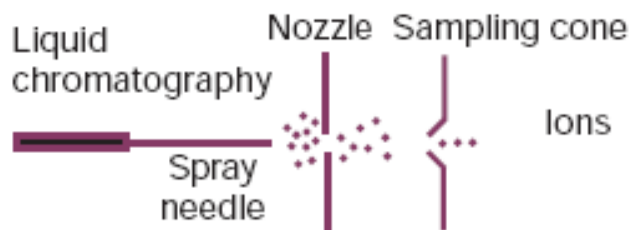


soft ion source →



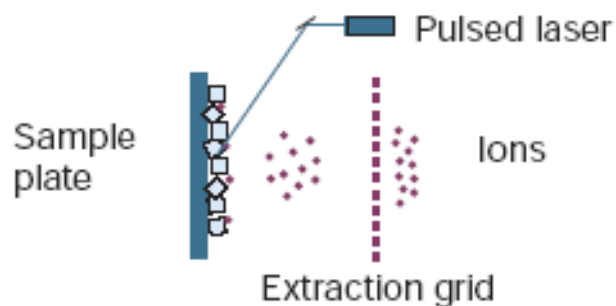
Ionization/ protonation

Soft ion source for bio-molecules



電灑游離化
Electrospray ionization (ESI)

低鹽類容忍度，帶多電荷，
能搭配液相層析分離

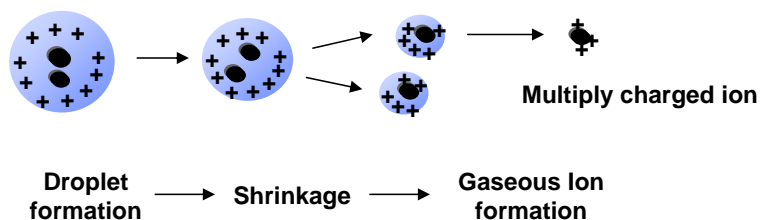
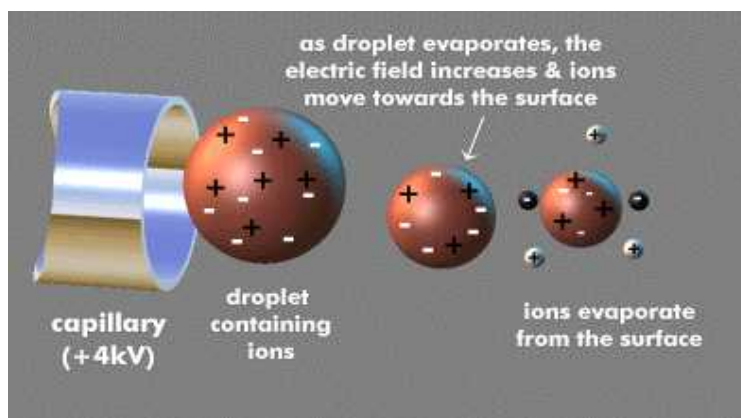


介質輔助雷射脫附游離化
Matrix-assisted laser desorption/ionization (MALDI)

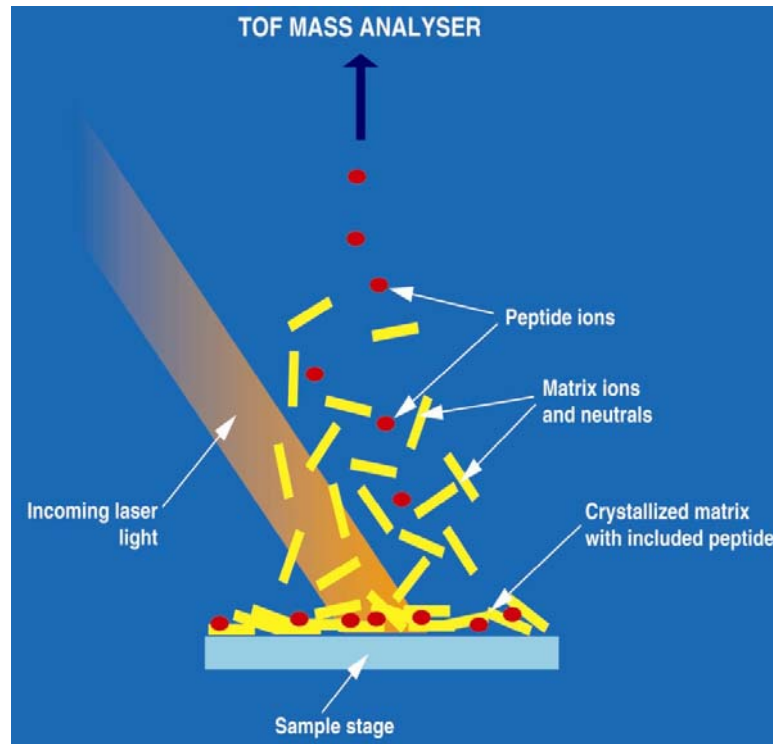
較高鹽類容忍度，帶單電荷

Nature 2003

Electrospray ionization (ESI)

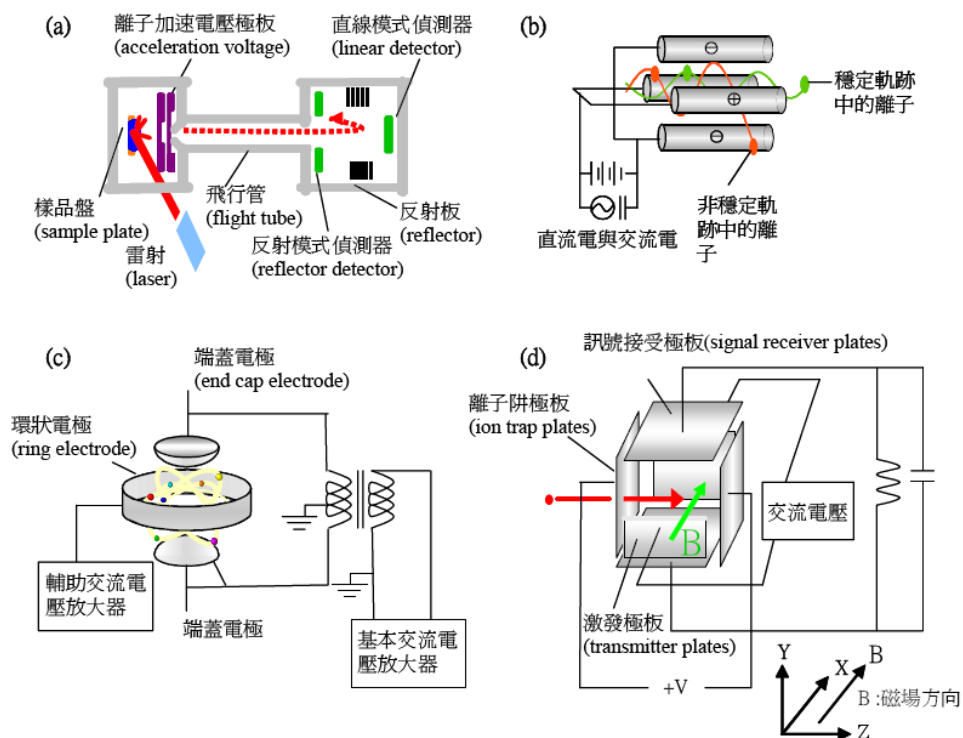


Matrix-assisted laser desorption ionization (MALDI)

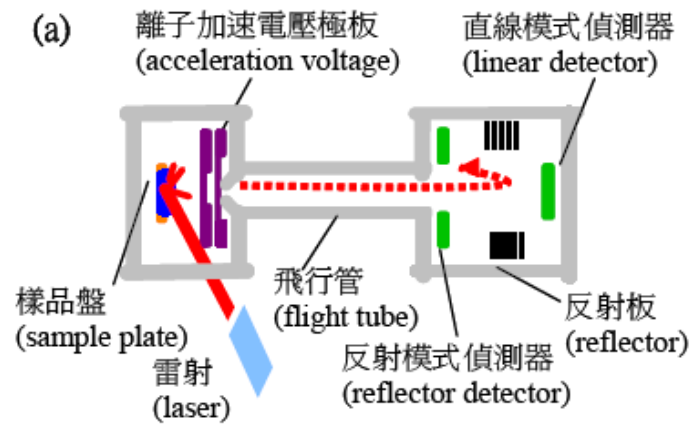


Mass analyzers

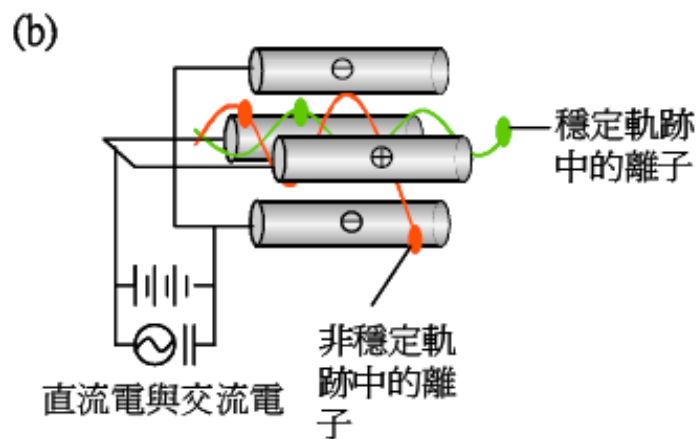
Four basic mass analyzers



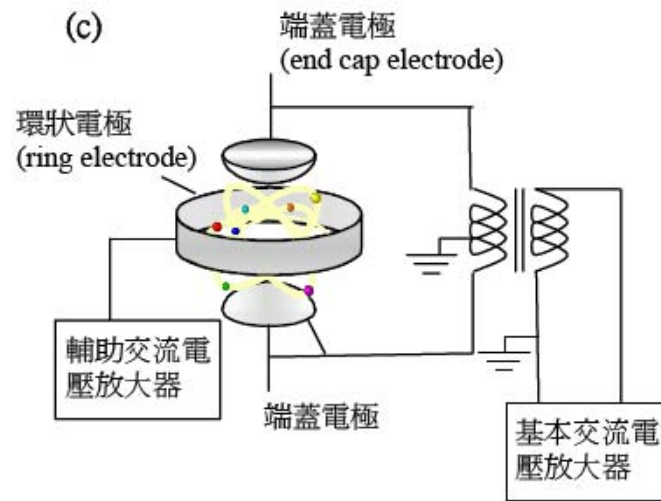
Time of flight (TOF)



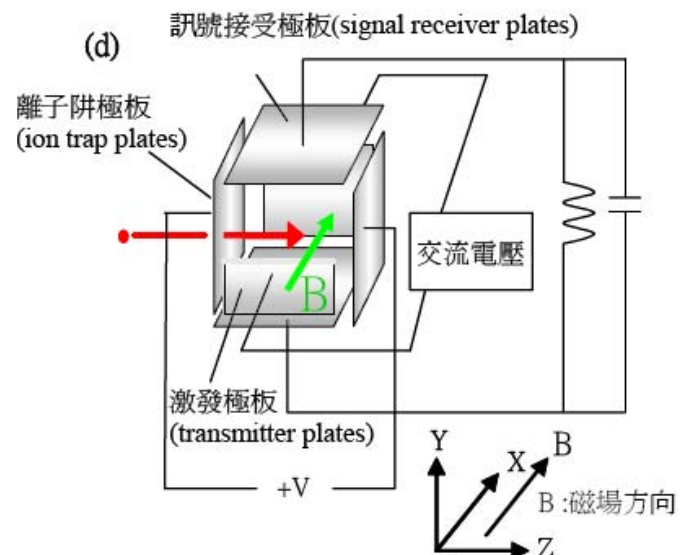
Quadrupole (Q)



Ion trap (IT)

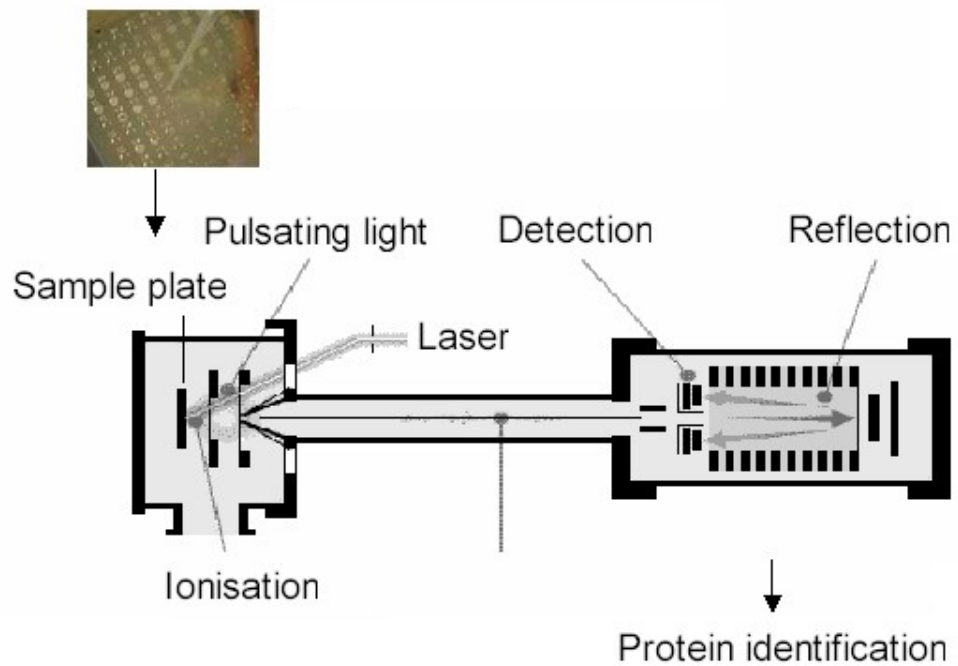


Fourier transform ion cyclotron resonance (FT-ICR)



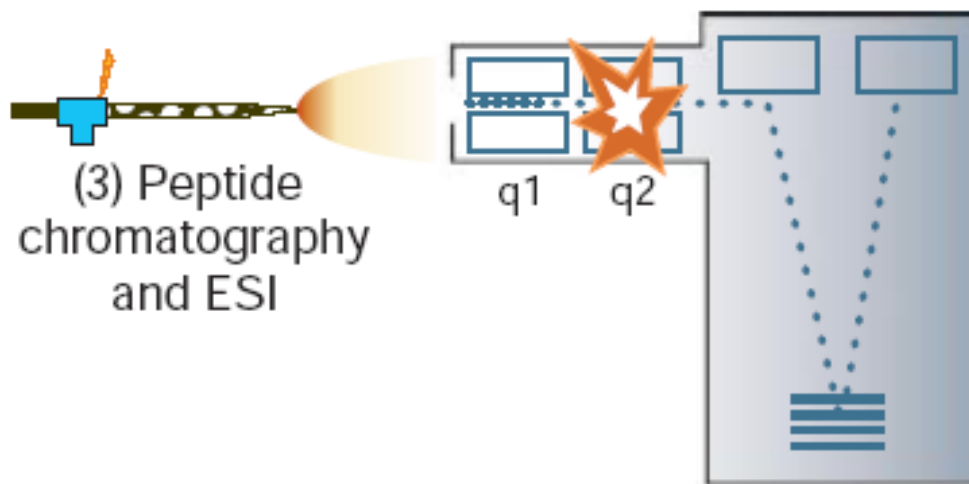
An example for the combination of ion source and mass analyzer

MALDI-TOF



An example for the combination of ion source and mass analyzer

ESI-Q/TOF



Tandem mass spectrometry

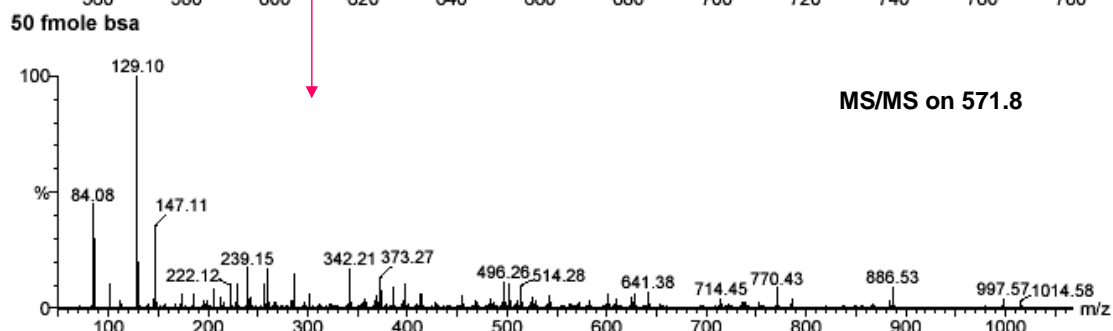
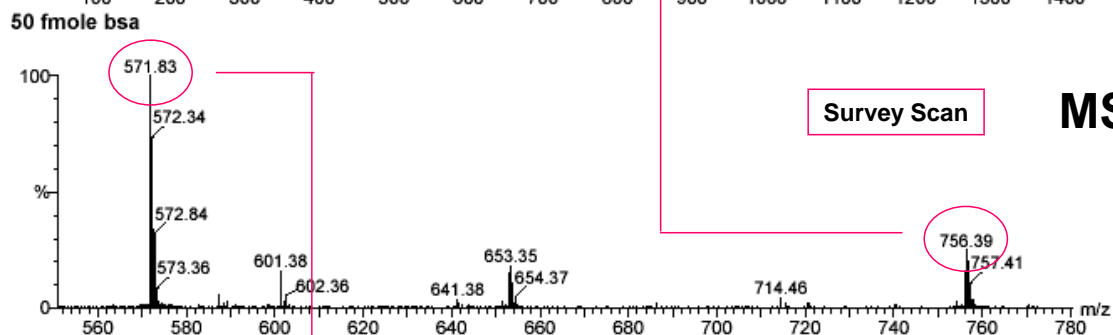
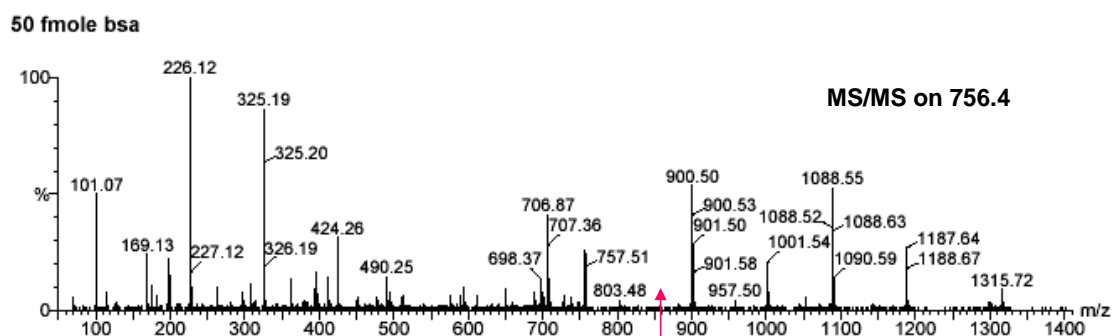
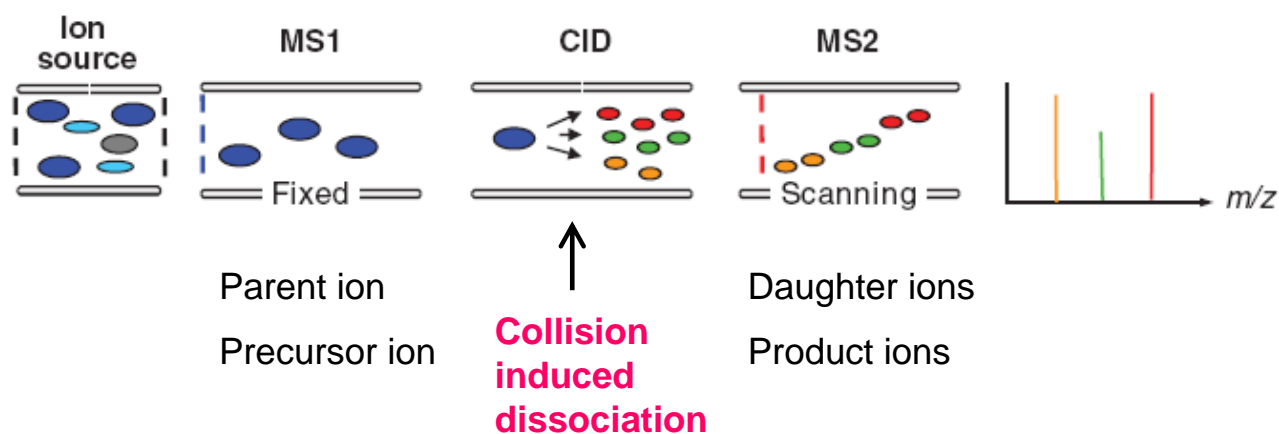
MS mode

MS/MS mode

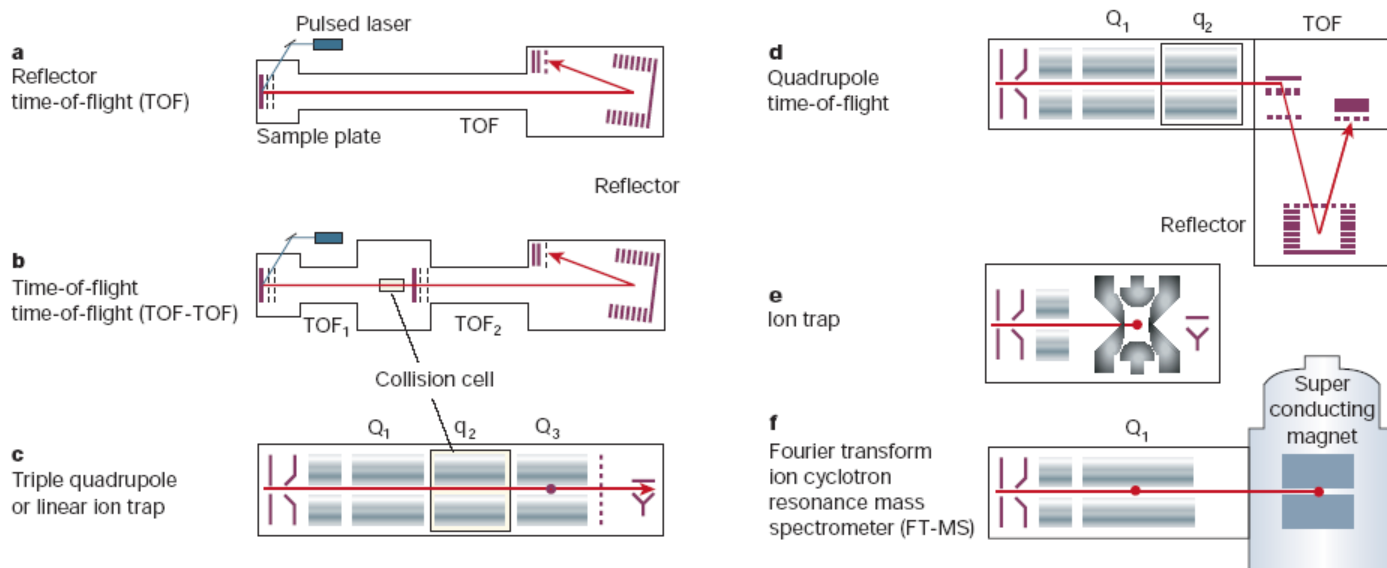
Parent ion MW

First MS: "select" the parent ion

Second MS: daughter ion detection



Some commercial mass spectrometers



And more...

General comparison

Table 1. Characteristics and performances of commonly used types of mass spectrometers. Check marks indicate available, check marks in parentheses indicate optional. +, ++, and +++ indicate possible or moderate, good or high, and excellent or very high, respectively. Seq., sequential.

	IT-LIT	Q-Q-ToF	ToF-ToF	FT-ICR	Q-Q-Q	QQ-LIT
Mass accuracy	Low	Good	Good	Excellent	Medium	Medium
Resolving power	Low	Good	High	Very high	Low	Low
Sensitivity (LOD)	Good		High	Medium	High	High
Dynamic range	Low	Medium	Medium	Medium	High	High
ESI	✓	✓		✓	✓	✓
MALDI	(✓)	(✓)	✓			
MS/MS capabilities	✓	✓	✓	✓	✓	✓
Additional capabilities	Seq. MS/MS			Precursor, Neutral loss, MRM		
Identification	++	++	++	+++	+	+
Quantification	+	+++	++	++	+++	+++
Throughput	+++	++	+++	++	++	++
Detection of modifications	+	+	+	+		+++

Protein Identification

1. Peptide mass mapping
2. LC-MS/MS

Protein digestion (bottom-up approach)

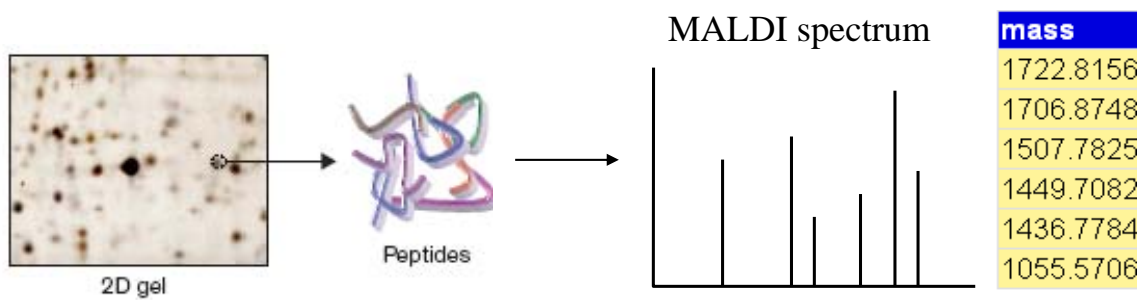
Sequence information					
Length: 147 AA [This is the length of the unprocessed precursor]	Molecular weight: 16135 Da [This is the MW of the unprocessed precursor]				
10	20	30	40	50	60
MRSLLILVLC	FLPLAALGKV	YGRCELAAAM	KRLGLDNYRG	YSLGNWVCAA	KFESNFNTHA
70	80	90	100	110	120
TNRNTDGSTD	YGILQINSRW	WCNDGRTPGS	KNLCNIPCSA	LLSSDITASV	NCAKKIASGG
130	140				
NGMNAWVAWR	NRCKGTDVHA	WIRGCR			



Trypsin digestion (enzyme specificity)

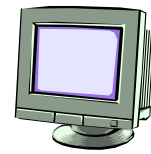
mass	position	#MC	modifications	peptide sequence
2337.1247	92-114	0		NLCNIPCSALLSSDITASVN CAK
1753.8351	64-79	0		NTDGSTDYGILQINSR
1589.7641	116-130	0		IASGGNGMNAWVAWR
1437.6505	52-63	0		FESNFNTHATNR
1268.6092	40-51	0		GYSLGNWVCAAK
1054.5428	135-143	0		GTDVHAWIR
936.3781	80-86	0		WWCNDGR
850.4417	33-39	0		LGLDNYR
836.4004	24-31	0		CELAAAMK

Peptide mass fingerprint (PMF)



Database search

Protein ID



```

10      20      30      40      50      60
                                     MQV
70      80      90      100     110     120
WPIGKk FE TLSYLPLTR dqlkeveyl lrk GWVPCLE FELEK gfvyr ehkspgyyd
130     140     150     160     170     180
arvwtm LP MFGTTDASOV LKELDEVVAA YPOAFV iia fdnvr OVOCI SFIAHTPESY
    
```

Characteristics for PMF

- Fast and convenient
- Purified compound is required (2DE + PMF)
- MW only. No sequence (fragments) information

MS/MS for peptide sequencing

Peptide: ^N **ABCDEF**G ^C

↓ MSMS

A **BCDEFG**

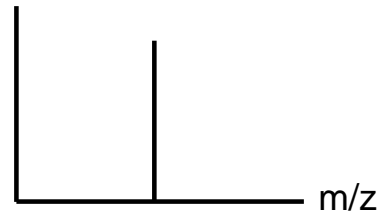
AB **CDEFG**

ABC **DEFG**

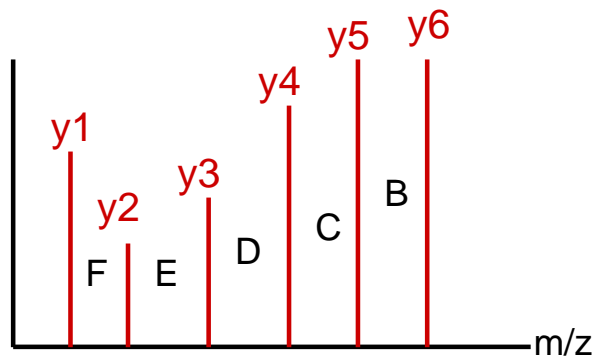
ABCDE **FG**

ABCDEF **G**

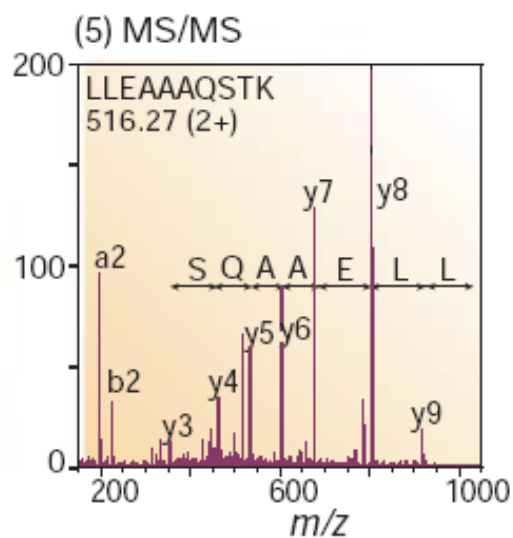
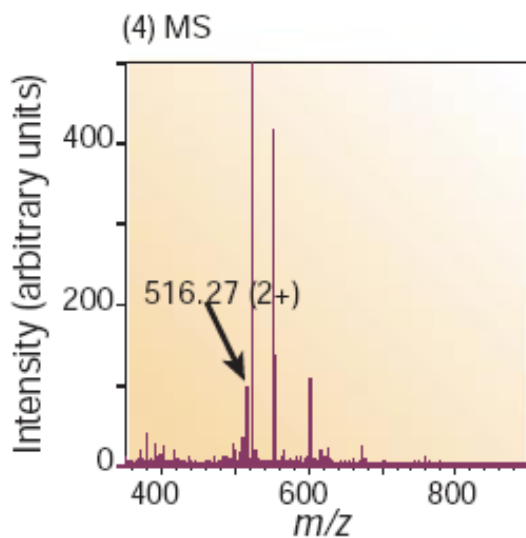
B_n ions y_n ions



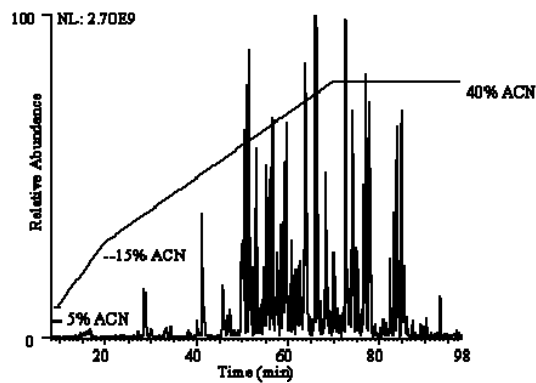
↓ MSMS



MS/MS for peptide sequencing

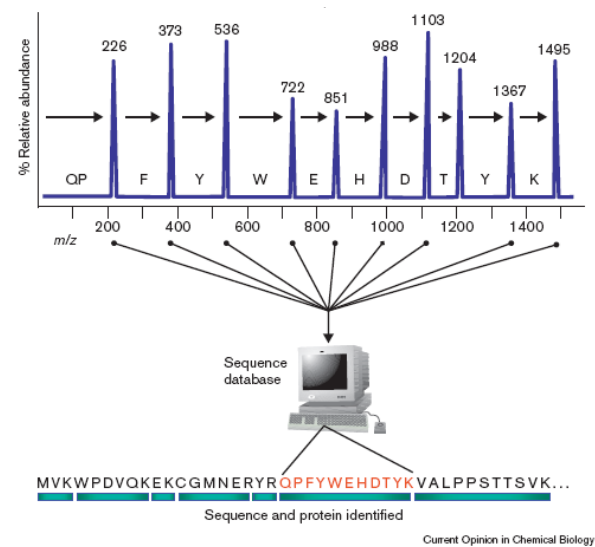


LC-MS/MS for protein identification



LC separation

High throughput protein ID



Current Opinion in Chemical Biology

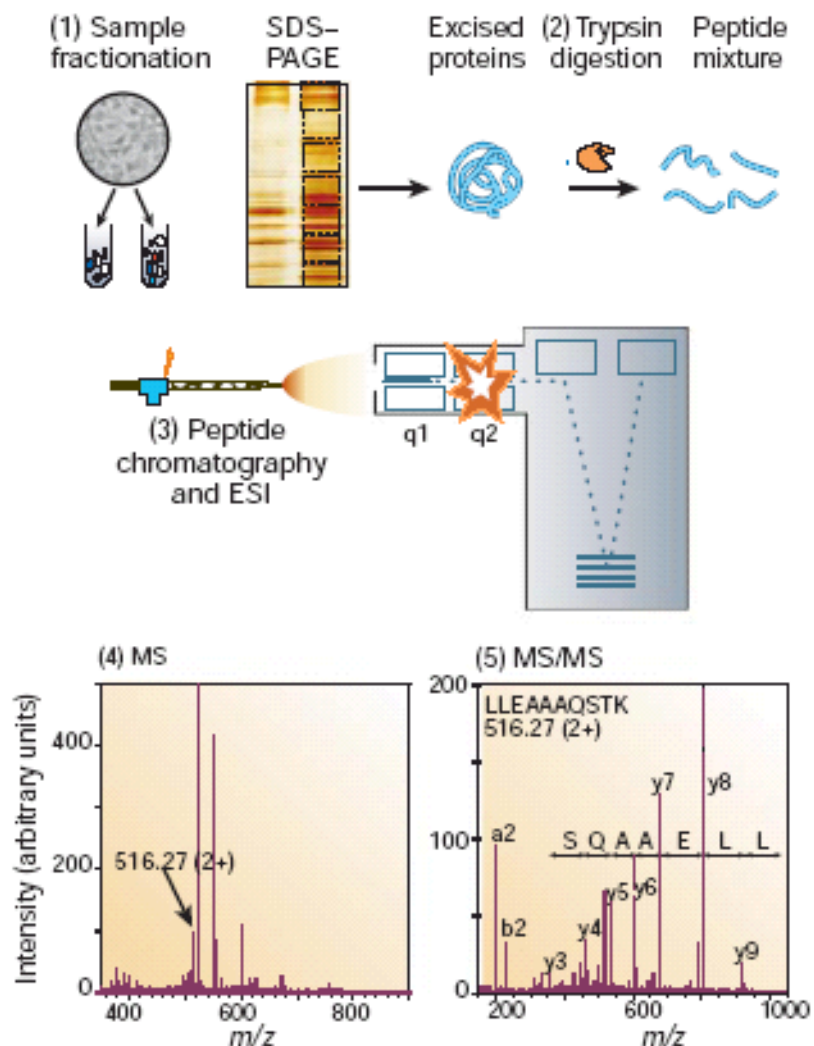
Characteristics for MS/MS based protein ID

- Sequence information
- Power to deal with complicated samples
- One peptide, one protein

LC-MSMS



Protein ID flowchart



Basic concept for routine protein analysis

Protein separation



Enzymatic digestion



Peptide separation



Mass spectrometry



Database search

Database search

MASCOT MS/MS Ions Search

Your name	<input type="text" value="Cindy"/>	Email	<input type="text" value="shengyu.huang@mass-solutior"/>
Search title	<input type="text"/>		
Database	<input type="text" value="SwissProt"/>		
Taxonomy	<input type="text" value="All entries"/>		
Enzyme	<input type="text" value="None"/>	Allow up to	<input type="text" value="1"/> missed cleavages
Fixed modifications	<input type="text" value="Biotin (K)
Biotin (N-term)
Carbamidomethyl (C)
Carbamyl (K)
Carbamyl (N-term)"/>	Variable modifications	<input type="text" value="Acetyl (K)
Acetyl (N-term)
Acetyl (Protein N-term)
Amidated (C-term)
Amidated (Protein C-term)"/>
Quantitation	<input type="text" value="None"/>		
Peptide tol. ±	<input type="text" value="0.5"/> Da	MS/MS tol. ±	<input type="text" value="0.5"/> Da
Peptide charge	<input type="text" value="1+, 2+ and 3+"/>	Monoisotopic	<input checked="" type="radio"/> Average <input type="radio"/>
Data file	<input type="text"/> <input type="button" value="浏览..."/>		
Data format	<input type="text" value="Micromass (.PKL)"/>	Precursor	<input type="text"/> m/z
Instrument	<input type="text" value="ESI-QUAD-TOF"/>	Error tolerant	<input type="checkbox"/>
Decoy	<input type="checkbox"/>	Report top	<input type="text" value="AUTO"/> hits
<input type="button" value="Start Search ..."/>		<input type="button" value="Reset Form"/>	

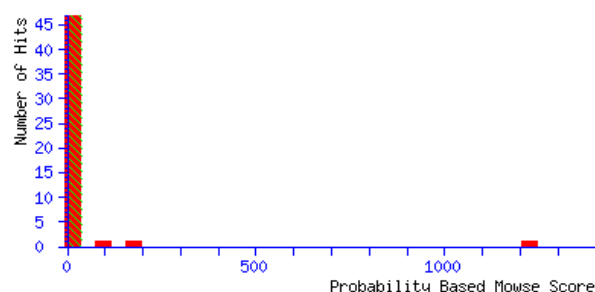
Mascot database search

{MATRIX} Mascot Search Results

User :
Email :
Search title : MS/MS Example
MS data file : C:\Auto MSMS output\Sample 1.pk1
Database : SwissProt 51.6 (257964 sequences; 93947433 residues)
Timestamp : 19 Feb 2007 at 14:09:47 GMT
Protein hits : [CH60_HUMAN](#) 60 kDa heat shock protein, mitochondrial precursor (Hsp60) (60 kDa c
[CH60_DROME](#) 60 kDa heat shock protein, mitochondrial precursor (Hsp60) (60 kDa c
[CH60_CAEEL](#) Chaperonin homolog Hsp-60, mitochondrial precursor (Heat shock prote
[CH60_XANAC](#) 60 kDa chaperonin (Protein Cpn60) (groEL protein) - Xanthomonas axor

Probability Based Mowse Score

Ions score is $-10 \cdot \log(P)$, where P is the probability that the observed match is a random event.
 Individual ions scores > 38 indicate identity or extensive homology ($p < 0.05$).
 Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



1. [CH60_HUMAN](#) Mass: 61016 Score: 1225 Queries matched: 31
 60 kDa heat shock protein, mitochondrial precursor (Hsp60) (60 kDa chaperonin) (CPN60) (Heat shock
 Check to include this hit in error tolerant search

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Peptide
<input checked="" type="checkbox"/> 11	417.1822	832.3498	832.3828	-0.0329	0	45	0.016	1	K.APGFGDNR.K
<input checked="" type="checkbox"/> 12	422.7433	843.4720	843.5066	-0.0346	0	46	0.017	1	K.VGEVIVTK.D
<input checked="" type="checkbox"/> 13	430.7328	859.4510	859.4837	-0.0327	0	36	0.15	1	K.IPAMTIAK.N + Oxidation (M)
<input checked="" type="checkbox"/> 15	451.2499	900.4853	900.5280	-0.0428	0	52	0.0039	1	K.LSDGVAVLK.V
<input checked="" type="checkbox"/> 16	456.7806	911.5467	911.5804	-0.0337	0	59	0.00056	1	K.VGLQVVAVK.A
<input checked="" type="checkbox"/> 21	480.7447	959.4748	959.5036	-0.0288	0	45	0.017	1	R.VTDALNATR.A
<input checked="" type="checkbox"/> 24	595.7855	1189.5565	1189.6012	-0.0447	0	(57)	0.0011	1	K.EIGNIISDAMK.K
<input checked="" type="checkbox"/> 25	603.7720	1205.5294	1205.5962	-0.0668	0	60	0.00048	1	K.EIGNIISDAMK.K + Oxidation (M)
<input checked="" type="checkbox"/> 26	608.3099	1214.6052	1214.6507	-0.0455	0	73	2.2e-05	1	K.NAGVEGSLIVEK.I
<input checked="" type="checkbox"/> 27	617.2857	1232.5569	1232.5885	-0.0316	0	81	4e-06	1	K.VGGTSDVEVNEK.K
<input checked="" type="checkbox"/> 31	672.8375	1343.6605	1343.7085	-0.0480	0	64	0.00016	1	R.TVIEQSWGSPK.V
<input checked="" type="checkbox"/> 34	714.8884	1427.7623	1427.8058	-0.0435	0	(65)	0.00014	1	R.GVMLAVDAVIAELK.K
<input checked="" type="checkbox"/> 35	714.8938	1427.7730	1427.8058	-0.0327	0	(73)	2.1e-05	1	R.GVMLAVDAVIAELK.K
<input checked="" type="checkbox"/> 36	722.8849	1443.7552	1443.8007	-0.0455	0	75	1.2e-05	1	R.GVMLAVDAVIAELK.K + Oxidation (M)
<input checked="" type="checkbox"/> 37	722.8934	1443.7722	1443.8007	-0.0285	0	(73)	2.2e-05	1	R.GVMLAVDAVIAELK.K + Oxidation (M)
<input checked="" type="checkbox"/> 39	752.8643	1503.7141	1503.7490	-0.0349	0	90	4.3e-07	1	K.TLNDELEIEGPK.F
<input checked="" type="checkbox"/> 40	760.8461	1519.6777	1519.7439	-0.0662	0	(89)	4.7e-07	1	K.TLNDELEIEGPK.F + Oxidation (M)
<input checked="" type="checkbox"/> 45	640.3281	1917.9625	1918.0636	-0.1010	0	102	2.1e-08	1	K.ISSIQSIIVPALEIANAHR.K
<input checked="" type="checkbox"/> 46	960.0327	1918.0509	1918.0636	-0.0127	0	(87)	5.1e-07	1	K.ISSIQSIIVPALEIANAHR.K
<input checked="" type="checkbox"/> 48	1019.5106	2037.0067	2037.0153	-0.0087	0	52	0.0015	1	R.IQEIEQLDVTTSEYK.E
<input checked="" type="checkbox"/> 51	1057.0537	2112.0929	2112.1323	-0.0394	0	116	6.8e-10	1	R.ALMLQGVLLADAVAVTMGPK.G
<input checked="" type="checkbox"/> 52	1065.0399	2128.0653	2128.1272	-0.0619	0	(72)	1.7e-05	1	R.ALMLQGVLLADAVAVTMGPK.G + Oxidati

Protein View

Match to: **ALBU_BOVIN** Score: 5041
Serum albumin precursor (Allergen Bos d 6) (BSA) - Bos taurus (Bovine)
Found in search of C:\Documents and Settings\MST\@à±\çà«Ø|''s\BSA.RAW

Nominal mass (M_r): 71244; Calculated pI value: 5.82

NCBI BLAST search of **ALBU_BOVIN** against nr

Unformatted [sequence string](#) for pasting into other applications

Taxonomy: [Bos taurus](#)

Fixed modifications: Carbamidomethyl (C)

Cleavage by Trypsin: cuts C-term side of KR unless next residue is P

Sequence Coverage: 77%

Matched peptides shown in **Bold Red**

```
1 MKWVTFISLL LLFSSAYSRG VFRRDTHKSE IAHRFKDLGE EHFKGLVLIA
51 FSQYLQCCPF DEHVKLVNEL TEFAKTCVAD ESHAGCEKSL HTLFGDELCK
101 VASLRETYGD MADCCEKQEP ERNECFLSHK DDSPDLPKLK PDPNTLCDEF
151 KADEKKFWGK YLYEIARRHP YFYAPELLY ANKYNGVFQE CCQAEDKGAC
201 LLPKIETMRE KVLASSARQR LRCASIQKFG ERALKAWSVA RLSQKFPKAE
251 FVEVTKLVTD LTKVHKECCH GDLLECADDR ADLAKYICDN QDTISSKLKE
301 CCDKPLLEKS HCIAEVEKDA IPENLPLTA DFAEDKDVCK NYQEAKDAFL
351 GSFLYESRR HPEYAVSVLL RLAKEYEATL EECCAKDDPH ACYSTVFDKL
401 KHLVDEPQNL IKQNCDQFEK LGEYGFQNAL IVRYTRKVPQ VSTPTLVEVS
451 RSLGKVGTRC CTKPESERMP CTEDYLSLIL NRLCVLHEKT PVSEKVTKCC
501 TESLVNRRRPC FSALTPDETY VPKAFDEKLF TFHADICTLP DTEKQIKKQT
551 ALVELLKHKP KATEEQLKTV MENFVAFVDK CCAADDKEAC FAVEGPKLVV
601 STQTALA
```

1. **ALBU_BOVIN** Mass: 71244

Serum albumin precursor (A

Query	Observed	Mr (expt)
2884	609.3470	608.339
3199	660.3850	659.377
3544	689.4460	688.438
3648	706.4700	705.462
3889	752.4310	751.423
3954	758.3630	757.355
4168	789.6450	788.637
4496	409.3366	816.658
4521	409.9631	817.911
5068	424.3146	846.614

Peptide View

MS/MS Fragmentation of **GACLLPK**

Found in **ALBU_BOVIN**, Serum albumin precursor (Allergen Bos d 6) (BSA) - Bos taurus (Bovine)

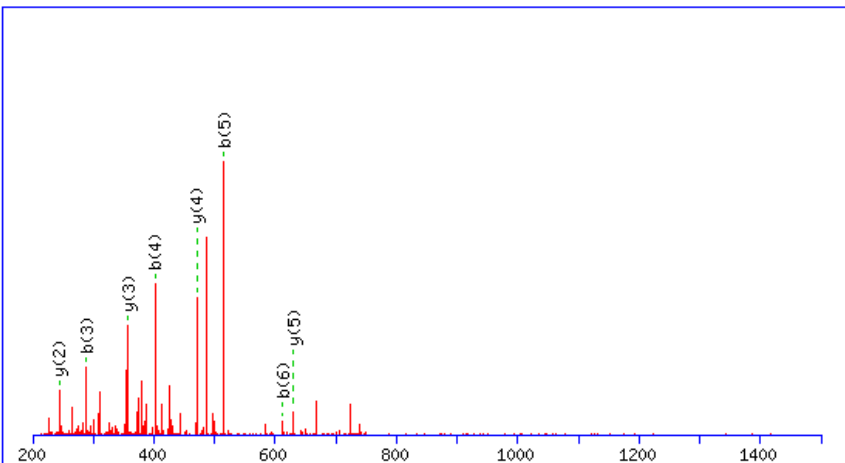
Match to Query 3954: 757.355724 from(758.363000,1+)

Title: BSA.7504.7504.1.dta

Data file C:\Documents and Settings\MST\@à±\çà«Ø|''s\BSA.RAW

Click mouse within plot area to zoom in by factor of two about that point

Or, to Da



1. **ALBU_BOVIN** Mass: 71244

Serum albumin precursor (A

Query	Observed	Mr (expt)
2884	609.3470	608.339
3199	660.3850	659.377
3544	689.4460	688.438
3648	706.4700	705.462
3889	752.4310	751.423
3954	758.3630	757.355
4168	789.6450	788.637
4496	409.3366	816.658
4521	409.9631	817.911
5068	424.3146	846.614

Perspective for MS based Proteomics

- Protein identification
- Post translational modification
- Function characterization
- Expression/Quantitation

Bioinformatics



Data processing/ comparison
protein database integration

Clinical use: drug screening, biomarker discovery etc.