MALDI-TOF/TOF MS 基础与应用

微生物研究所

孙 树 涛

2011-07-21

内容概要

- 质谱技术发展回顾
- **MALDI-TOF/TOF MS构造与原理**
- 分析策略和实验流程
- 样品前处理
- 数据检索和结果分析
- **MALDI-TOF/TOF MS应用**

质谱技术发展回顾

- **1906**年，**J.J.Thompson**因发明质谱技术并用以研究气体的电导获得诺贝尔物理奖；
- **1922**年，**F.W.Aston**因用质谱仪发现了非放射性元素同位素获得诺贝尔化学奖；
- **1980**年，**W.Paul**因发明离子阱质谱原理和技术获得诺贝尔物理奖；
- **1996**年，**R.E.Smally**等因用质谱发现 **C60**获得诺贝尔化学奖；
- **2002**年，**J.B.Fenn**，**K.Tanaka**因发明生物大分子质谱电离技术获得诺贝尔化学奖

质谱

样品在离子源内电离，产生的各种离子经过质量分析器按其质荷比 **(m/z)** 分离，并依次被检测器检测并记录下来，形成一个按顺序记录各种质荷比（**m/z**）离子相对丰度的谱图。

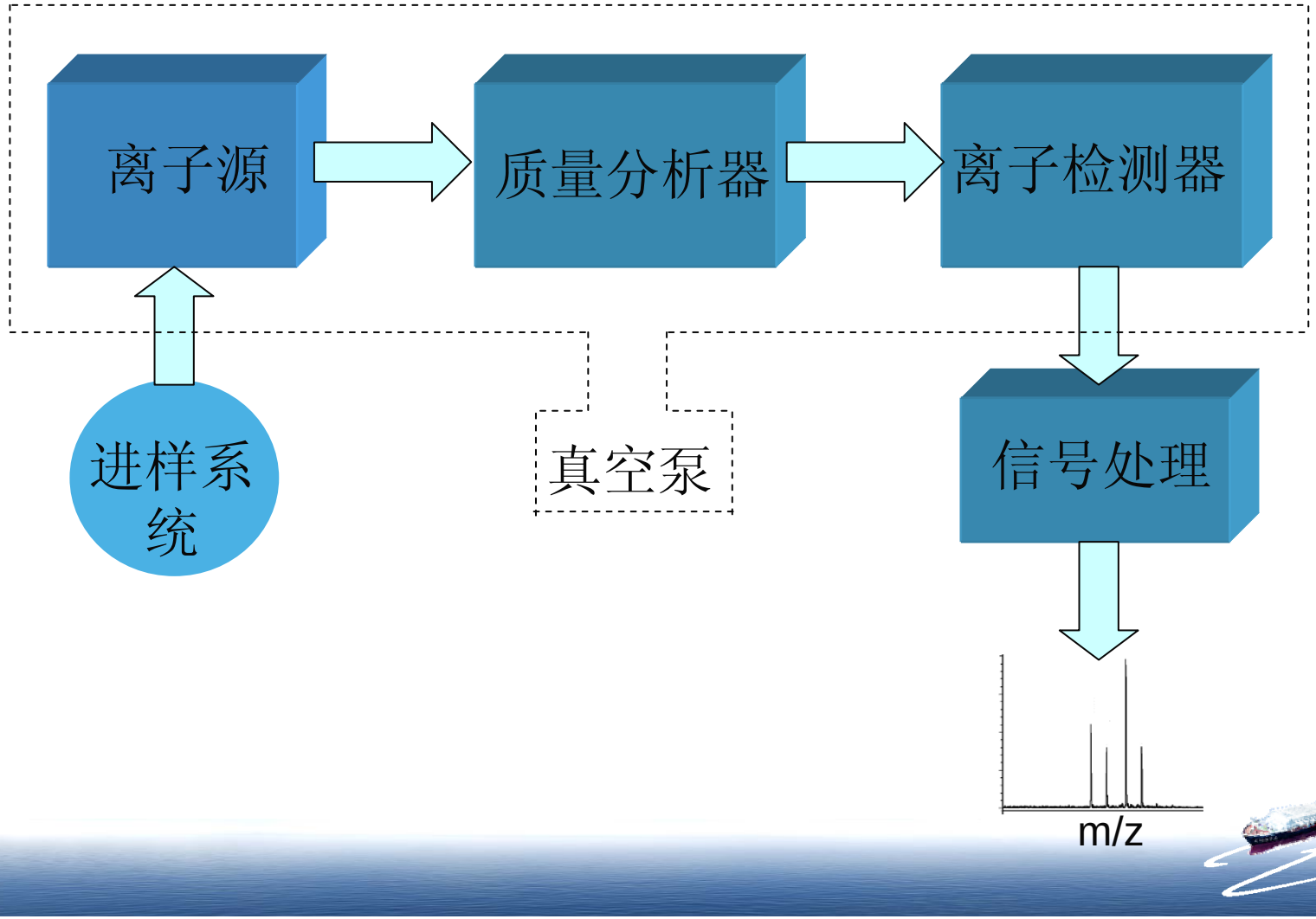
特点

- **Sensitivity**
- **Speed**
- **Specificity**
- **Stoichiometry**

灵敏度的比较

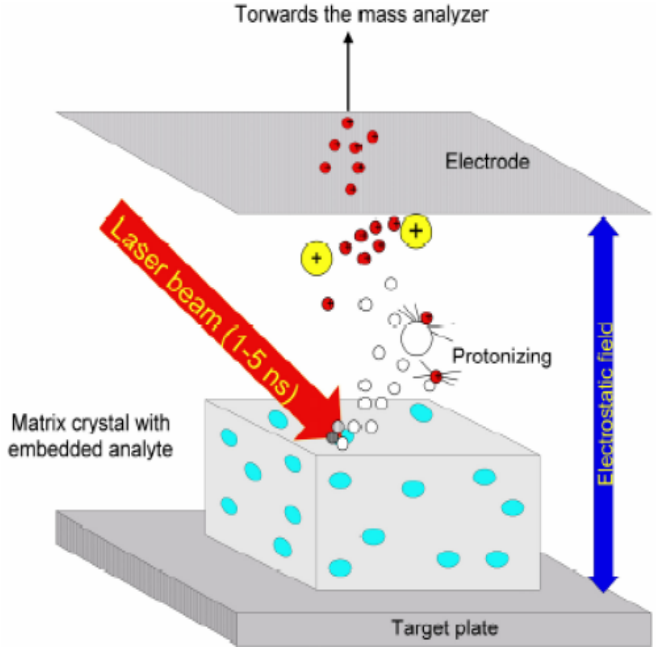
分析方法	检测灵敏度（g）
NMR	10^{-3}～10^{-5}
IR	10^{-6}～10^{-7}
GC	10^{-6}～10^{-13}
UV	10^{-6}～10^{-7}
MS	10^{-11}～10^{-12}

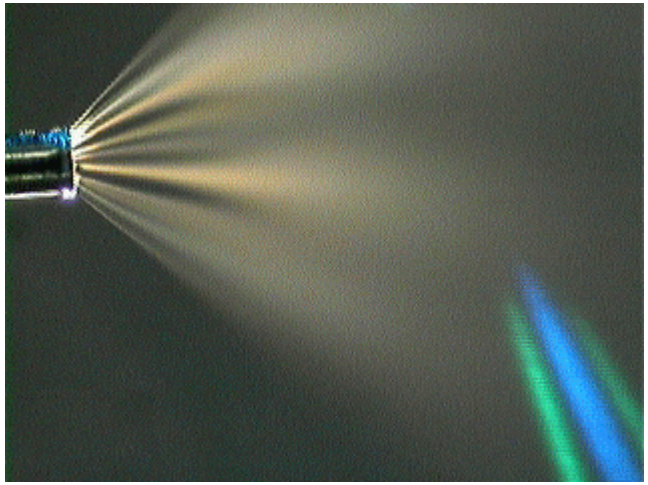
质谱仪器基本构造



离子源

生物大分子的电离





- MALDI

  Matrix-assisted laser desorption/ionization

  基质辅助激光解吸离子化

- ESI

  Electrospray ionization

  电喷雾电离

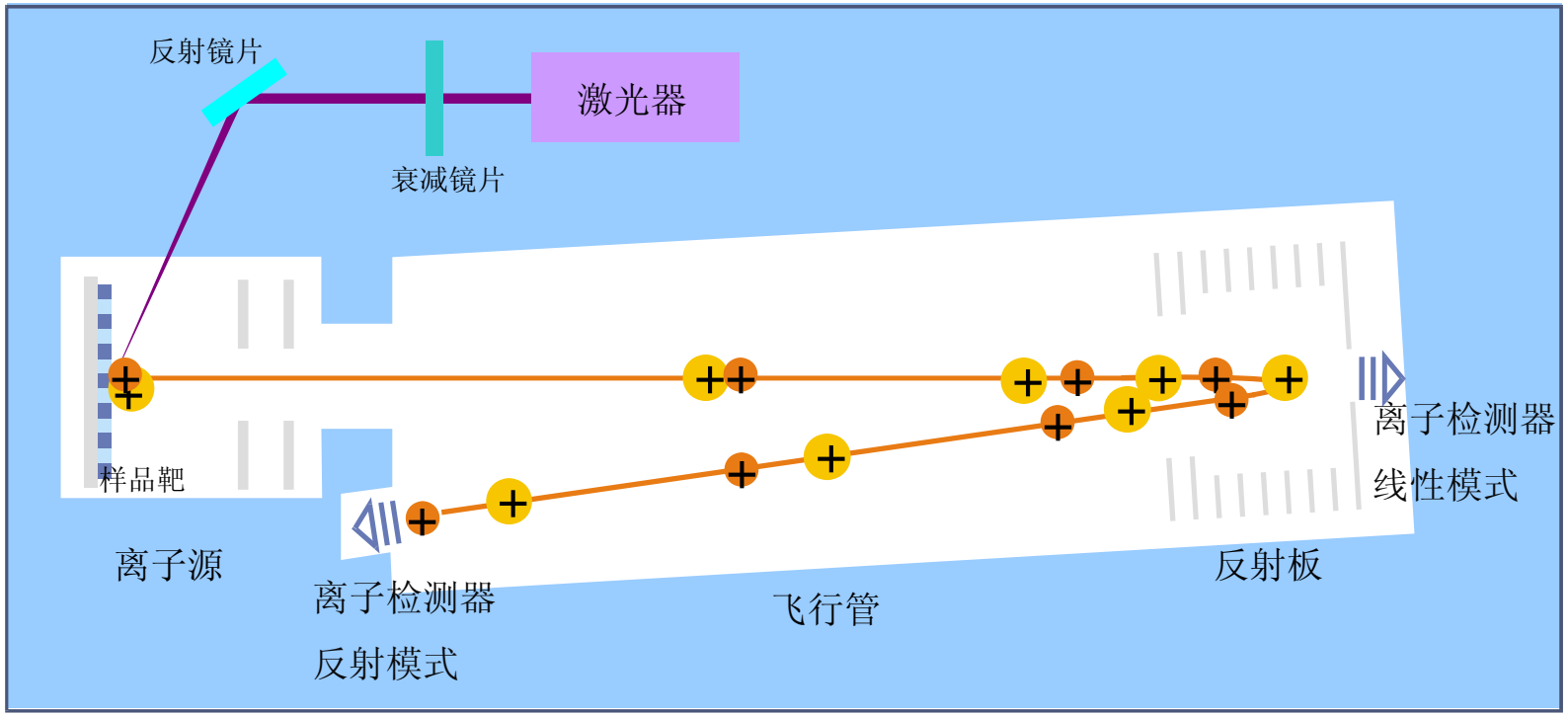
质量分析器

- 电磁扇型双聚焦质谱仪**(Magnetic Sector)**
- 飞行时间质谱仪 **(TOF, time of flight)**
- 四极杆质谱仪 **(Quadrupole)**
- 离子阱质谱仪 **(Ion trap)**
- 傅立叶变换离子回旋共振质谱仪 **(FT ICR)**
- 静电场轨道阱质谱仪 **(Orbitrap)**

飞行时间质谱

TOF (time of flight)

线性模式(linear mode)，分辨率较低；反射模式(reflectron mode)，分辨率高

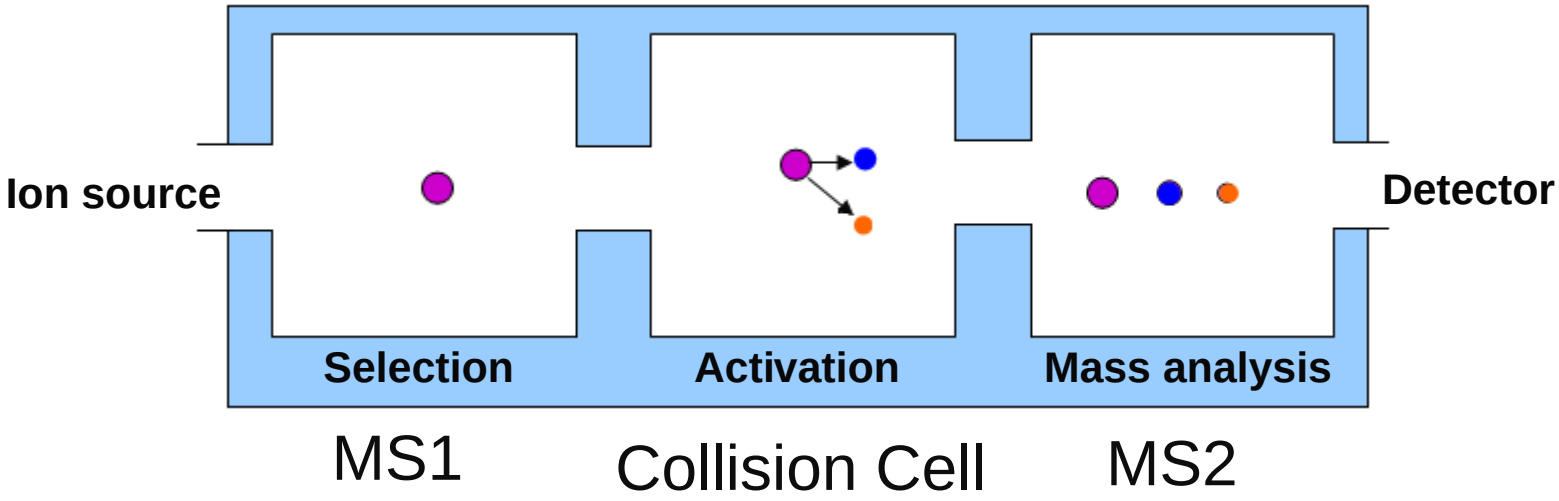


$KE = qV = 1/2\ mv^2 \rightarrow v = (2qV/m)^{1/2} \rightarrow t = L/v = L / (2V/m/q)^{1/2}$

$\rightarrow m/z = m/q = 2V\ t^2 / L^2$

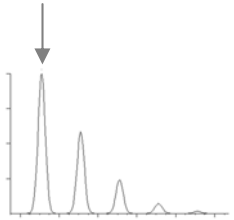
串级质谱（MS/MS)

串联质谱是通过对选定某个**m/z**母离子反应产物的分析，对目标化合物进行断裂机理或结构鉴定等方面的研究。



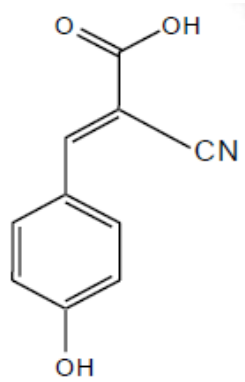
基础概念

- 分辨率（Resolution） $R = \frac{M}{\Delta M}$
- 信噪比（S/N）
- 灵敏度（Sensitivity） 最低检测限fmol,amol
- 质量精度（Accuracy） $\left| \frac{M - M_{。}}{m} \right|$ ppm
- 稳定性（Stability）



- 单同位素质量(Monoisotopic mass)
- 质荷比（m/z） $\frac{M + z}{z}$

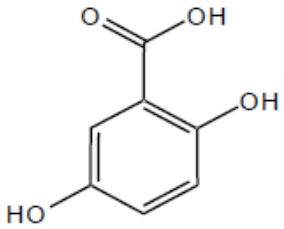
常用基质



a-Cyano-4-hydroxycinnamic acid

CHCA

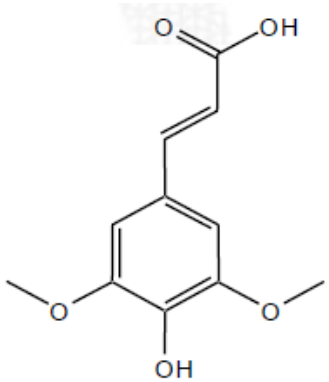
a－氰基－4－羟基肉桂酸



2,5-Dihydroxybenzoic acid

DHB

2，5－二羟基苯甲酸



Sinapinic acid

SA

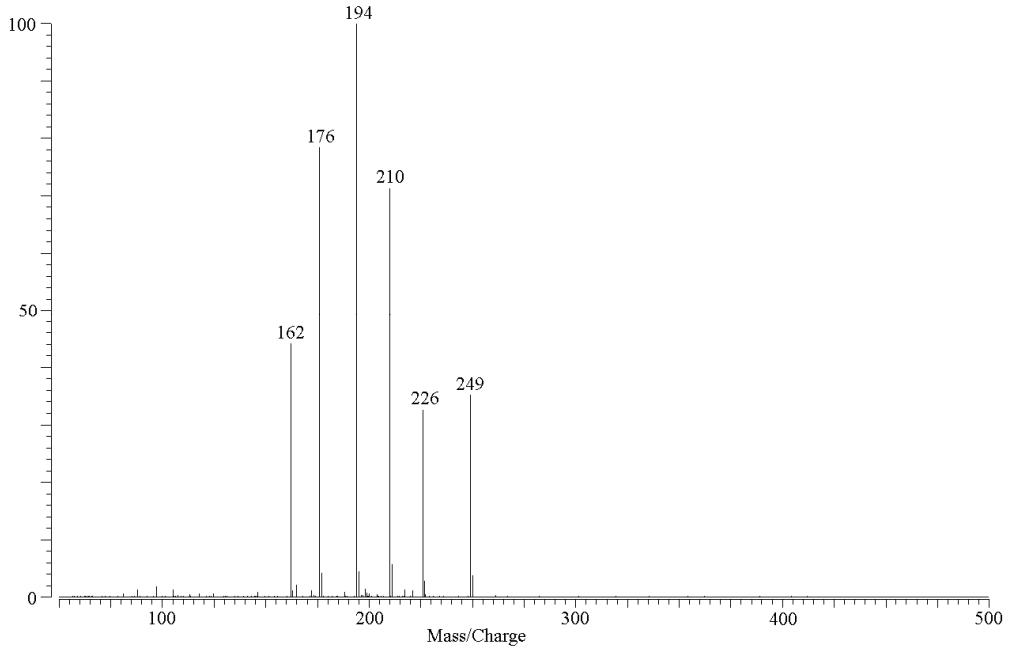
芥子酸

CHCA主要用于20KDa以下的蛋白质

DHB和SA可用于20KDa以上的蛋白质

低质量区 (<700Da) 基质背景信号干扰

纳米材料作为基质



七种氨基酸混合溶液的MALDI质谱

MALDI-TOF/TOF MS 特点

- 可得到MS/MS谱，得到肽段氨基酸组成信息
- 分辨率高　～20000
- 灵敏度高　～fmol,amol
- 分析速度快，通量高
- 质量精度高 ～20ppm
- 分子量测量范围大 ～200000Da
- 对样品中的盐等杂质有一定的容忍性

MALDI-TOF/TOF MS 的应用范围

- 分子量测定

  蛋白、多肽、脂类、糖类、寡聚核苷酸、天然产物等

- 蛋白质鉴定

  1D, 2D-GE、单一蛋白溶液酶解液

- 蛋白质翻译后修饰(PTM)研究

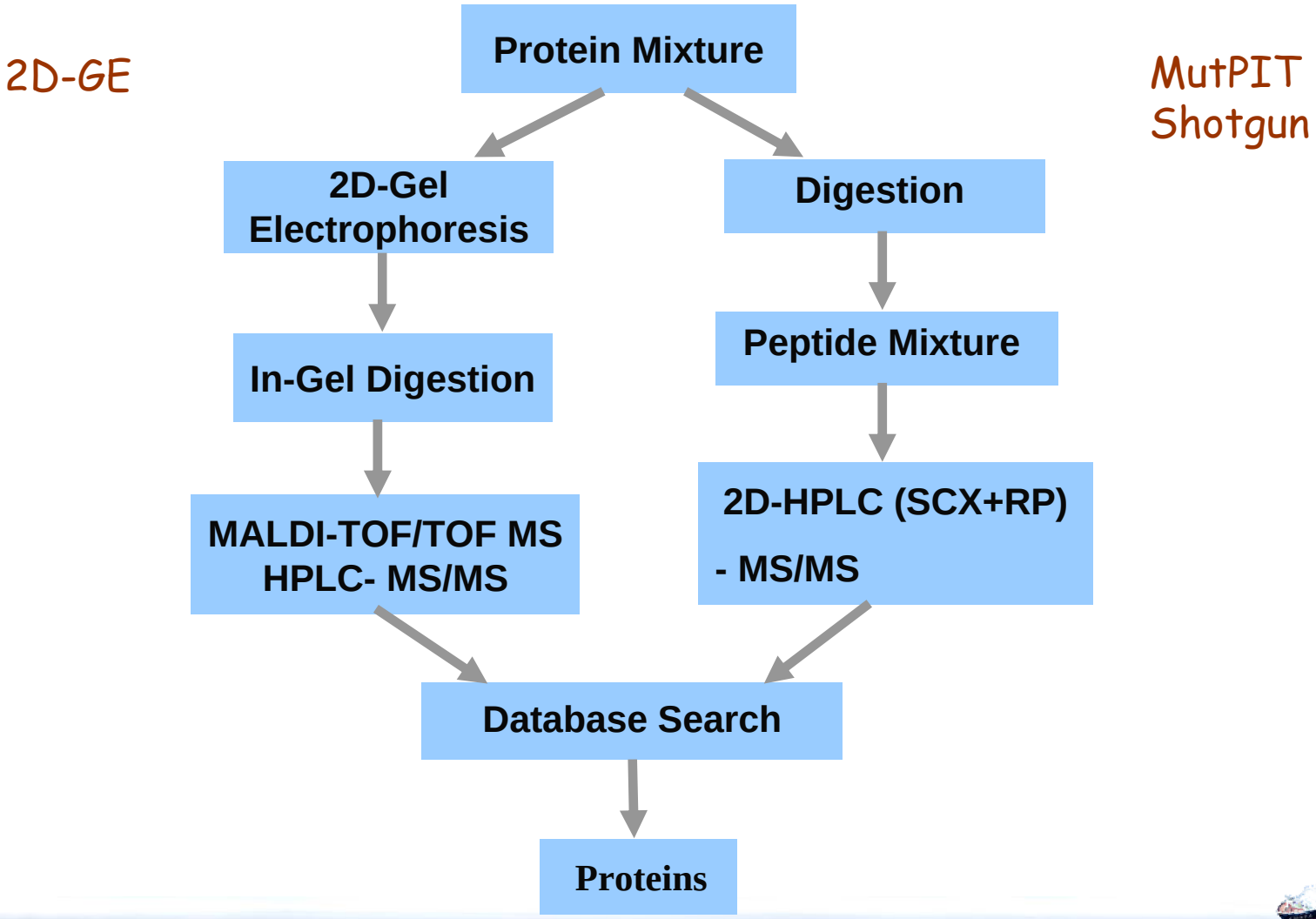
- 疾病诊断Biomarker筛选

基于质谱的蛋白质组分析策略

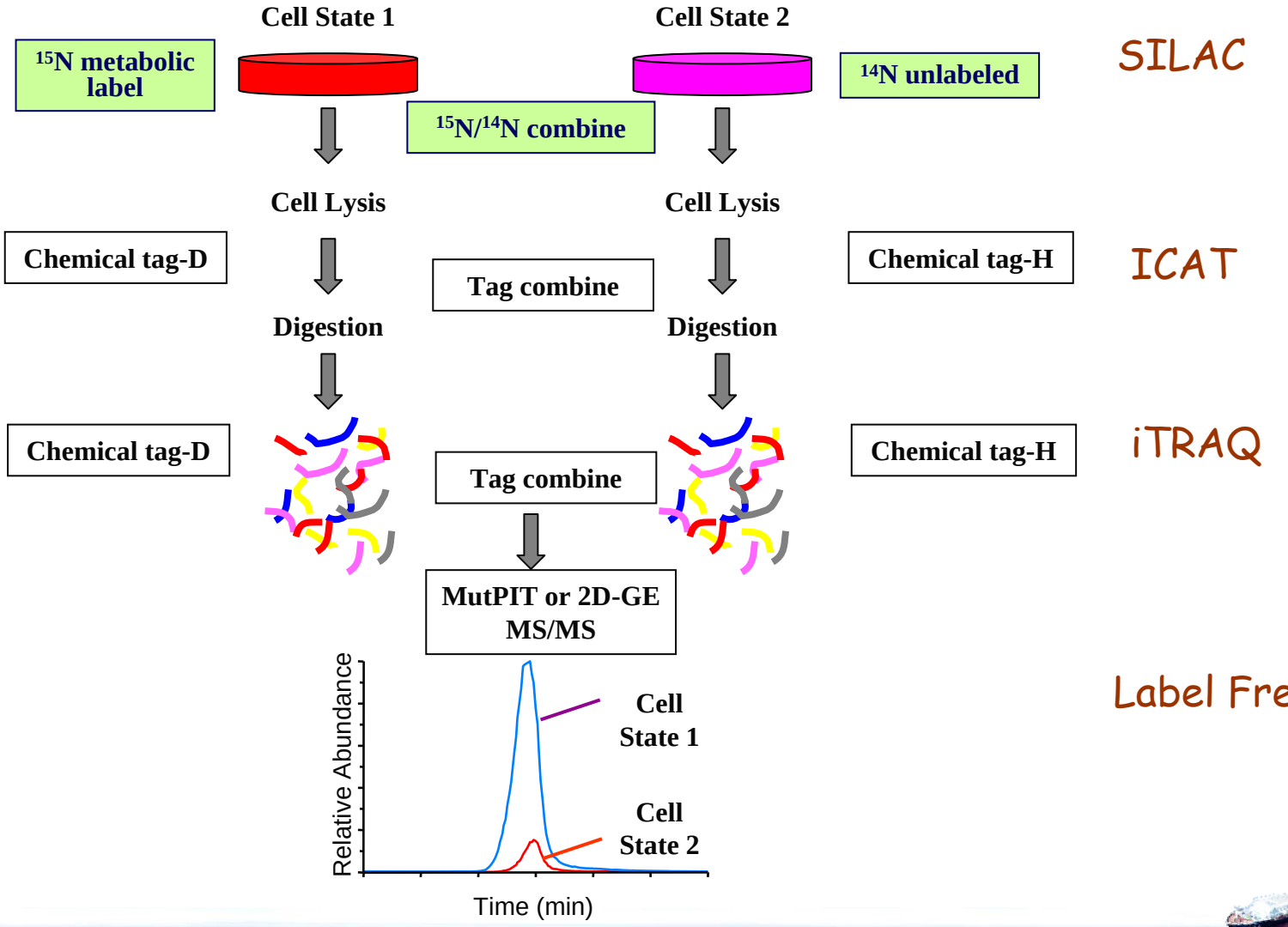
获取样品 → 样品前处理（提取、分离） → 质谱鉴定 → 数据检索、评估、整合 → 得到结果

• Cells • Tissues • Plasma • Urine	• Protein extraction • Protein separation • Digestion • Peptide separation • PTM enrichment	• TOF • TOF-TOF • Ion trap • FT-ICR • Orbitrap	•Databases •MASCOT •SEQUEST	•Identification •Quantification •Function

基于质谱的蛋白质组学实验流程



定量


Cell State 1
Cell State 2
15N metabolic label
14N unlabeled
15N/14N combine
SILAC
Cell Lysis
Cell Lysis
Chemical tag-D
Chemical tag-H
Tag combine
ICAT
Digestion
Digestion
Chemical tag-D
Chemical tag-H
Tag combine
iTRAQ
MutPIT or 2D-GE MS/MS
Relative Abundance
Cell State 1
Cell State 2
Time (min)
Label Free

样品处理注意事项

- 用进口的离心管
- 防止皮肤、头发的角蛋白（戴无粉手套），及BSA污染
- 制备过程中防止发生修饰和降解
- 水的纯度：二次去离子水，电阻率≥18MΩ·cm
- 不希望有钠、钾加合离子
- 常用溶剂：水、乙氰、乙醇、甲醇、丙酮、甲酸、三氟乙酸等
- 避免使用SDS、Triton、CHAPS等去垢剂, PEG等聚合物对质谱信号造成严重干扰的物质

样品分子量测定

- 待分析的蛋白样品处于溶解状态或冻干粉，要求不含盐，不含去垢剂
- 提供样品浓度为几十 pm/μL，一次提供样品10μL左右即可
- 送样要求：检测目的，样品来源、制备方法、溶剂、可能混有成分、分子量范围

蛋白鉴定

- 单一蛋白，考染能见的条带
- 需把蛋白溶液或胶内酶切成700-3500Da左右的肽段
- 染色的样品需先清洗、脱色、还原、烷基化、酶切
- 送样要求：蛋白物种来源、检测目的、制备方法、溶剂、可能混有成分

MALDI-TOF MS可容忍的最高浓度

Urea	0.5M
Guanidine-HCl	0.5M
Glycerol	1%
Alkali metal salts	0.02M
Tris buffer	0.05M
NH_4HCO_3	0.05M
Phosphate buffer	0.01M
SDS, CHAPS, PEG, Triton	需去除

去除杂质干扰，浓缩样品

- 超滤
- 较多量样品过固相萃取柱（**SPE**）
- 微量样品过**ZipTip**柱
  - **C18** 去除**peptide,protein(<50KDa),oligonucletides**

    样品中盐的干扰及浓缩
  - **C4** 蛋白质**(20KDa-100KDa)**除盐及浓缩
  - **SCX** 去除样品中去垢剂及浓缩样品

数据采集

- 点靶
- 记录样品点在靶上位置A1-H24
- 选择采集方法
- 采集数据
- 导出数据
- 搜索数据库
- 分析数据

蛋白质鉴定

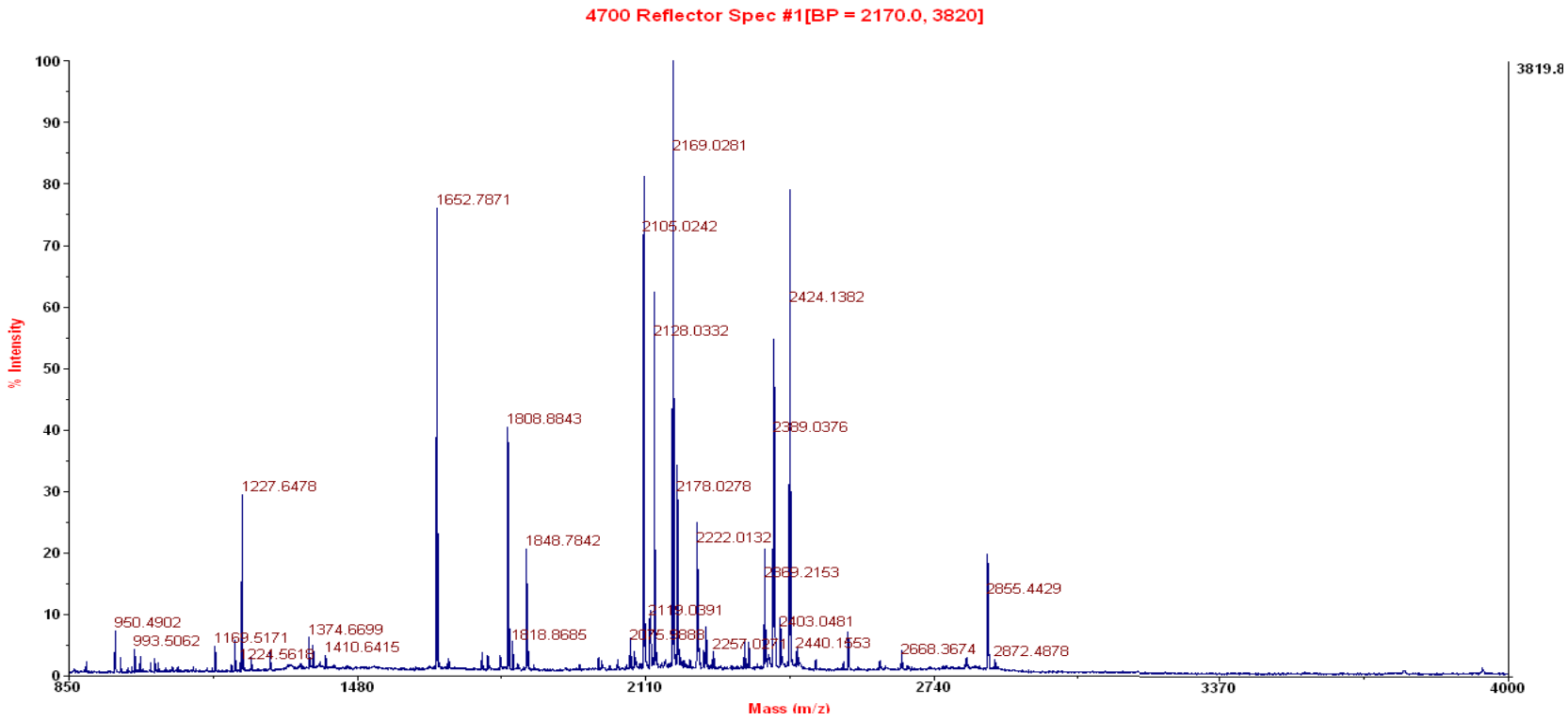
- 指纹谱（PMF, peptide mass fingerprint）

  蛋白质经酶切位点专一的蛋白酶水解后得到肽段混合物，经过质谱分析得到的质谱图。

  由于组成不同蛋白质的氨基酸序列不同，酶解后产生的肽段也不相同，其肽段混合物质量数具有特征性，所以称肽质量指纹谱。

- 优点：是蛋白质鉴定的经典方法，算法简单，速度快， 在串联质谱鉴定出现之前应用广泛。
- 缺点：质量相近的多肽增加匹配难度，并且无法实现混合蛋白的鉴定，不太适合数据库不完整的物种的蛋白质鉴定，不能分析到翻译后修饰位点

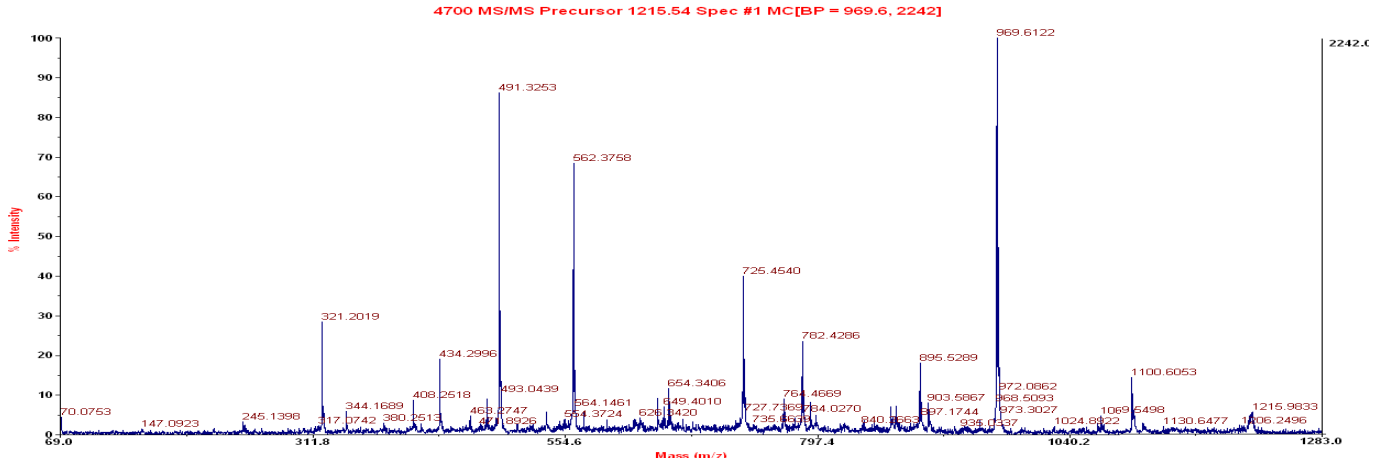
PMF谱图示例

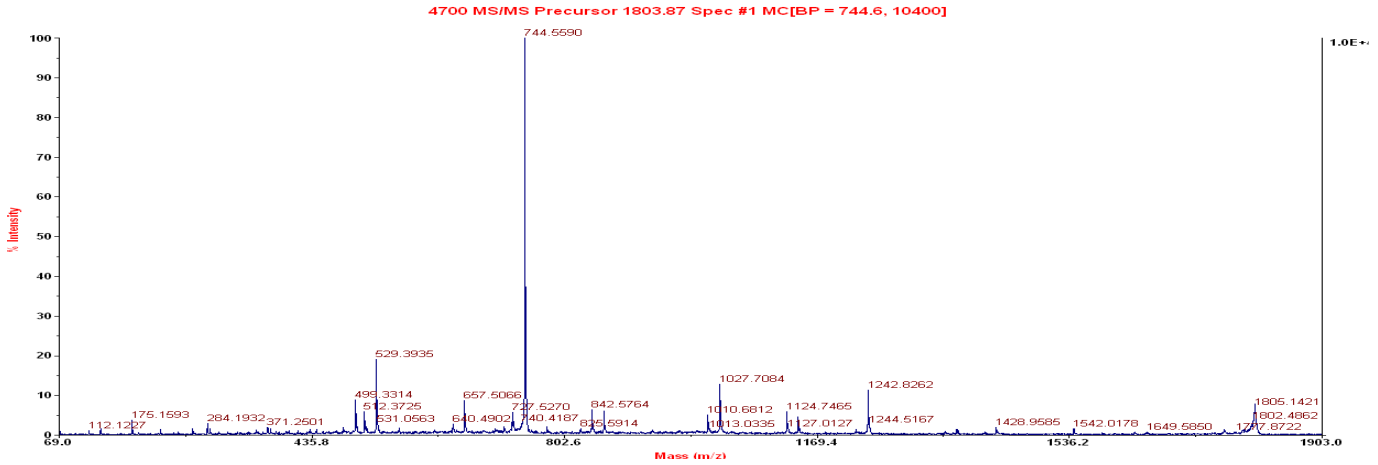

4700 Reflector Spec #1[BP = 2170.0, 3820]
% Intensity
Mass (m/z)
100
90
80
70
60
50
40
30
20
10
0
850
1480
2110
2740
3370
4000
3819.8
2169.0281
1652.7871
2105.0242
2424.1382
2128.0332
1808.8843
2389.0376
1227.6478
2178.0278
1848.7842
2222.0132
2855.4429
950.4902
993.5062
1374.6699
1410.6415
1618.8685
2403.0481
2668.3674
2872.4878

二级谱图（MS/MS）

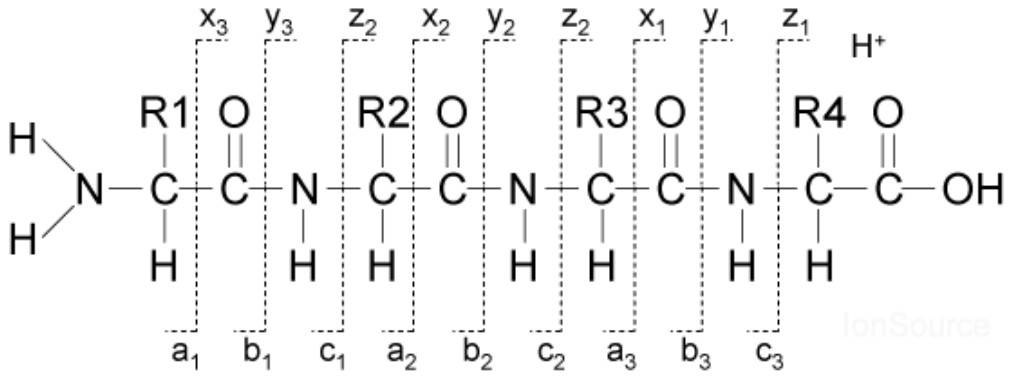
- 可提供结构组成信息，可用于对蛋白分离效果较差的样品。
- 原理：对一个或者多个未被解析的肽段MS/MS 数据进行对数据库比较从而进行蛋白鉴定。
- 优点：是目前应用最广的高通量鉴定蛋白质方法，鉴定准确度更高，无需人工序列解析，可以实现混合蛋白的鉴定。
- 缺点：增加了一步操作，对仪器要求更严格。

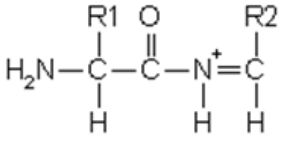
MS/MS谱图示例

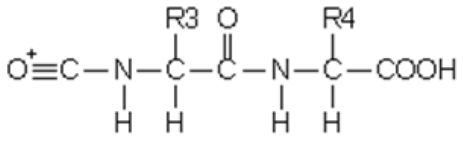

4700 MS/MS Precursor 1215.54 Spec #1 MC[BP = 969.6, 2242]
% Intensity
Mass (m/z)
969.6122
491.3253
562.3758
725.4540
321.2019
782.4286
434.2996
895.5289
1100.6053
654.3406
493.0439
408.2518
344.1689
70.0753
147.0923
245.1398
380.2513
1215.9833
69.0
311.8
554.6
797.4
1040.2
1283.0

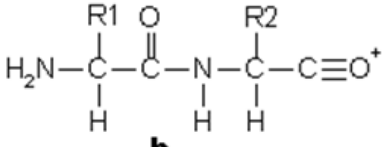

4700 MS/MS Precursor 1803.87 Spec #1 MC[BP = 744.6, 10400]
% Intensity
Mass (m/z)
744.5590
529.3935
1027.7084
1242.8262
499.3314
657.5066
842.5764
1124.7465
175.1593
284.1932
371.2501
1805.1421
1802.4862
69.0
435.8
802.6
1169.4
1536.2
1903.0

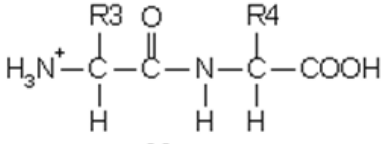
MS/MS肽段断裂的碎片命名规则



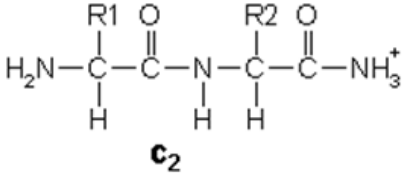


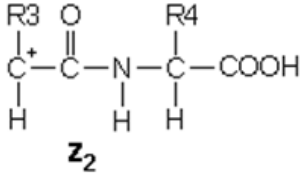






CID





ETD

Ion Type	Neutral M_r
a	[N]+[M]-CHO
a*	a-NH_3
a°	a-H_2O
b	[N]+[M]-H
b*	b-NH_3
b°	b-H_2O
c	[N]+[M]+NH_2
d	a - partial side chain
v	y - complete side chain
w	z - partial side chain
x	[C]+[M]+CO-H
y	[C]+[M]+H
y*	y-NH_3
y°	y-H_2O
z	[C]+[M]-NH_2

数据库检索

- Mascot　　http://www.matrixscience.com
  Local Server
- PMF数据库搜索时，将数据库中的蛋白序列理论酶切成肽段，计算其理论图谱，与实验图谱进行比对得到结果
- 扣除基质本底、酶自切和角蛋白污染峰

http://cn.expasy.org/tools/peptide-mass.html

理论酶切
选取酶（700-3500Da）
理论数据与实验谱图进行比对

PMF检索界面

MASCOT Peptide Mass Fingerprint

Your name	***	**Email**	***@**.***
Search title			
Database(s)	SwissProt NCBInr contaminants cRAP MSDB	**Enzyme**	Trypsin
		Allow up to	1 missed cleavages
Taxonomy	All entries		
Fixed modifications	Carbamidomethyl (C)	> <	Acetyl (K) Acetyl (N-term) Acetyl (Protein N-term) Amidated (C-term) Amidated (Protein C-term) Ammonia-loss (N-term C) Biotin (K) Biotin (N-term) Carbamyl (K) Carbamyl (N-term) Carboxymethyl (C)
	Display all modifications ☐		
Variable modifications	Oxidation (M)	> <	
Protein mass	kDa	**Peptide tol. ±**	100 ppm
Mass values	◉ MH^+ ○ M_r ○ $M\text{-}H^-$	**Monoisotopic**	◉ **Average** ○
Data file	浏览...		
Query NB Contents of this field are ignored if a data file is specified.			
Decoy	☐	**Report top**	AUTO hits
	Start Search ...		**Reset Form**

MS/MS检索界面

MASCOT MS/MS Ions Search

Your name	***	**Email**	***@***.***
Search title			
Database(s)	Fungi_EST Environmental_EST SwissProt NCBInr contaminants	**Enzyme**	Trypsin
		Allow up to	1 missed cleavages
		Quantitation	None
Taxonomy	All entries		
Fixed modifications	Carbamidomethyl (C)	> / <	mTRAQ (N-term) mTRAQ (Y) mTRAQ:13C(3)15N(1) (K) mTRAQ:13C(3)15N(1) (N-term) mTRAQ:13C(3)15N(1) (Y) NIPCAM (C) Oxidation (HW) Phospho (ST) Phospho (Y) Propionamide (C) Pyridylethyl (C)
	Display all modifications ☐		
Variable modifications	Oxidation (M)	> / <	
Peptide tol. ±	100 ppm **# ^{13}C** 0	**MS/MS tol. ±**	0.5 Da
Peptide charge	1+	**Monoisotopic**	◉ **Average** ○
Data file	浏览...		
Data format	Mascot generic	**Precursor**	m/z
Instrument	MALDI-TOF-TOF	**Error tolerant**	☐
Decoy	☐	**Report top**	AUTO hits
	Start Search ...		**Reset Form**

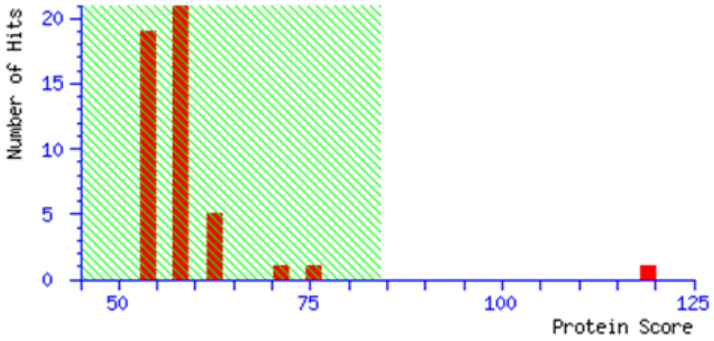
PMF检索结果示例

Mascot Search Results

```
User            : sst
Email           : shutaosun@im.ac.cn
Search title    : Project: 20100126, Spot Set: 20100126\12, Label: D4, Spot Id: 202120, Peak List Id: 441368, MS Job Run Id: 21426
MS data file    : pmf_D4_130376206501.txt
Database        : NCBInr 20110419 (13767831 sequences; 4728199773 residues)
Timestamp       : 25 Apr 2011 at 07:39:50 GMT
Top Score       : 119 for gi|12666724, glucoamylase [Talaromyces emersonii]
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 84 are significant (p<0.05).



Concise Protein Summary Report

Format As	Concise Protein Summary	Help
	Significance threshold p< 0.05	Max. number of hits AUTO

(continued)

Re-Search All | Search Unmatched

1. gi|12666724 Mass: 65902 Score: 119 Expect: 1.7e-05 Matches: 13
glucoamylase [Talaromyces emersonii]

2. gi|326932060 Mass: 57375 Score: 75 Expect: 0.48 Matches: 9
PREDICTED: tyrosine-protein kinase HCK-like [Meleagris gallopavo]

Search Parameters

Type of search	: Peptide Mass Fingerprint
Enzyme	: Trypsin
Fixed modifications	: Carbamidomethyl (C)
Variable modifications	: Oxidation (M)
Mass values	: Monoisotopic
Protein Mass	: Unrestricted
Peptide Mass Tolerance	: ± 100 ppm
Peptide Charge State	: 1+
Max Missed Cleavages	: 1
Number of queries	: 71
Selected for scoring	: 29

Mascot: http://www.matrixscience.com/

Mascot

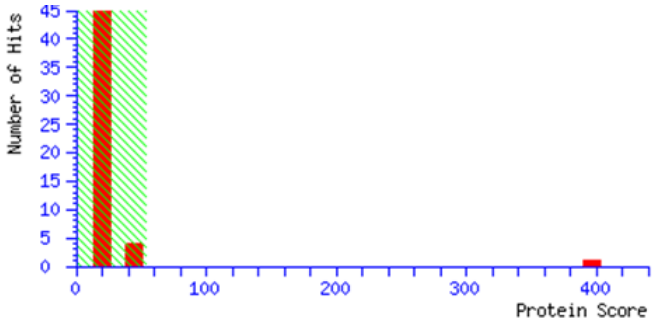
PMF分数算法

基于概率，显著性分析

- **S = -10*Log(P),**
- **P：比对匹配的一个随机事件的可能性大小**
- **P=E/N，E=期望值，**

  N=数据库中蛋白质数目的大小

MS/MS检索结果示例



Peptide Summary Report

Format As	Peptide Summary		Help
	Significance threshold p< 0.05	Max. number of hits AUTO	
	Standard scoring ◉ MudPIT scoring ○	Ions score or expect cut-off 0	Show sub-sets 0
	Show pop-ups ◉ Suppress pop-ups ○	Sort unassigned Decreasing Score	Require bold red ☐

Select All | Select None | Search Selected | ☐ **Error tolerant**

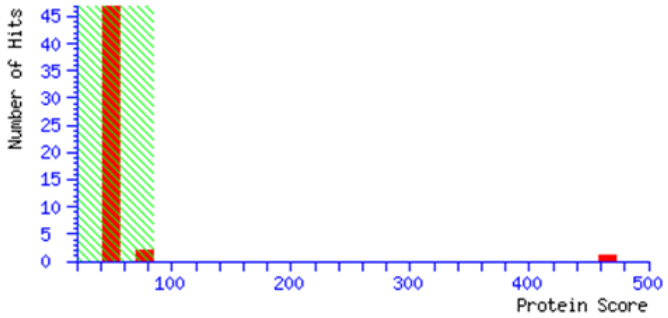
1. gi|12666724 **Mass:** 65902 **Score:** 396 **Matches:** 6(5) **Sequences:** 6(5)
glucoamylase [Talaromyces emersonii]
☐ Check to include this hit in error tolerant search

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Unique	Peptide
☑ 13	1435.5845	1434.5772	1434.6204	-30.11	0	47	0.19	1	U	R.SDPNYFYSWTR.D
☑ 23	1587.6808	1586.6735	1586.7213	-30.09	0	58	0.021	1	U	K.YTPSDGSLTEQFSR.T
☑ 38	1735.8329	1734.8256	1734.8788	-30.65	0	78	0.00024	1	U	K.DLEQTIQQYISAQAK.V
☑ 46	1834.9109	1833.9036	1833.9585	-29.89	0	68	0.0025	1	U	R.SIYAINSGIAEGSAVAVGR.Y
☑ 65	2440.1653	2439.1580	2439.2281	-28.75	0	70	0.0011	1	U	R.TDGTPLSASALTWSYASLLTASAR.R
☑ 76	3124.3113	3123.3040	3123.3663	-19.92	0	76	0.00014	1	U	K.DVNSILGSIHTFDPAGGCDDSTFQPCSAR.A

```
User            : sst
Email           : shutaosun@im.ac.cn
Search title    : Project: 20100126, Spot Set: 20100126\12, Label: D3, Spot Id: 202119, Peak List Id: 441373, MS Job Run Id: 21427
MS data file    : ppw_D3_130376210200.txt
Database        : NCBInr 20110419 (13767831 sequences; 4728199773 residues)
Timestamp       : 25 Apr 2011 at 07:44:29 GMT
Warning         : A Peptide summary report will usually give a much clearer picture of MS/MS search results.
Top Score       : 466 for gi|12666724, glucoamylase [Talaromyces emersonii]
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 84 are significant (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Protein Summary Report

Format As	Protein Summary (deprecated)		Help
	Significance threshold p< 0.05	Max. number of hits AUTO	

Index

	Accession	Mass	Score	Description
1.	gi|12666724	65902	466	glucoamylase [Talaromyces emersonii]
2.	gi|194859648	106930	67	GG23965 [Drosophila erecta]

Results List

1. gi|12666724 **Mass:** 65902 **Score:** 466 **Expect:** 3.5e-40 **Matches:** 18

glucoamylase [Talaromyces emersonii]

Observed	Mr(expt)	Mr(calc)	ppm	Start	End	Miss	Ions	Peptide
1114.5933	1113.5860	1113.6142	-25.30	223	233	0	---	R.ALVEGNALATR.L
1210.6034	1209.5961	1209.6394	-35.74	91	101	0	---	K.YLVDAFIAGNK.D
1435.5845	1434.5772	1434.6204	-30.11	73	83	0	---	R.SDPNYFYSWTR.D
1435.5845	1434.5772	1434.6204	-30.11	73	83	0	47	R.SDPNYFYSWTR.D
1587.6808	1586.6735	1586.7213	-30.09	420	433	0	---	K.YTPSDGSLTEQFSR.T
1587.6808	1586.6735	1586.7213	-30.09	420	433	0	58	K.YTPSDGSLTEQFSR.T
1735.8329	1734.8256	1734.8788	-30.65	102	116	0	---	K.DLEQTIQQYISAQAK.V
1735.8329	1734.8256	1734.8788	-30.65	102	116	0	78	K.DLEQTIQQYISAQAK.V
1759.7350	1758.7277	1758.7809	-30.24	584	598	0	---	K.NQTDGTIVWEDDPNR.S
1834.9109	1833.9036	1833.9585	-29.89	316	334	0	---	R.SIYAINSGIAEGSAVAVGR.Y
1834.9109	1833.9036	1833.9585	-29.89	316	334	0	68	R.SIYAINSGIAEGSAVAVGR.Y
1956.9371	1955.9298	1955.9800	-25.66	117	136	0	---	K.VQTISNPSGDLSTGGLGEPK.F
2440.1653	2439.1580	2439.2281	-28.75	434	457	0	---	R.TDGTPLSASALTWSYASLLTASAR.R
2440.1653	2439.1580	2439.2281	-28.75	434	457	0	70	R.TDGTPLSASALTWSYASLLTASAR.R
2927.4656	2926.4583	2926.5076	-16.84	91	116	1	---	K.YLVDAFIAGNKDLEQTIQQYISAQAK.V
2947.3777	2946.3704	2946.4076	-12.61	555	580	0	---	R.ADAYTNSNPLWYVTVNLPPGTSFEYK.F
3124.3113	3123.3040	3123.3663	-19.92	274	302	0	---	K.DVNSILGSIHTFDPAGGCDDSTFQPCSAR.A
3124.3113	3123.3040	3123.3663	-19.92	274	302	0	76	K.DVNSILGSIHTFDPAGGCDDSTFQPCSAR.A

No match to: 989.4664, 989.9616, 1025.4775, 1233.5402, 1236.5833, 1245.5348, 1302.5693, 1391.5854, 1434.5857, 1467.5940, 1477.5947, 1552.7001, 1571.7124, 1579.6885, 1595.7002, 1596.6962, 1596.6962, 1602.6752, 1612.6891,

Protein View

Match to: **gi|12666724** Score: **466** Expect: **3.5e-40**
glucoamylase [Talaromyces emersonii]
Found in search of ppw_D3_130376210200.txt

Nominal mass (M_r): **65902**; Calculated pI value: **4.44**
NCBI BLAST search of gi|12666724 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Talaromyces emersonii

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: 31%

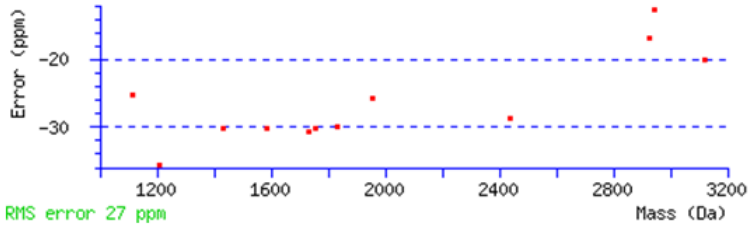
Matched peptides shown in **Bold Red**

```
    1 MASLVAGALC ILGLTPAAFA RAPVAARATG SLDSFLATET PIALQGVLNN
   51 IGPNGADVAG ASAGIVVASP SRSDPNYFYS WTRDAALTAK YLVDAFIAGN
  101 KDLEQTIQQY ISAQAKVQTI SNPSGDLSTG GLGEPKFNVN ETAFTGPWGR
  151 PQRDGPALRA TALIAYANYL IDNGEASTAD EIIWPIVQND LSYITQYWNS
  201 STFDLWEEVE GSSFFTTAVQ HRALVEGNAL ATRLNHTCSN CVSQAPQVLC
  251 FLQSYWTGSY VLANFGGSGR SGKDVNSILG SIHTFDPAGG CDDSTFQPCS
  301 ARALANHKVV TDSFRSIYAI NSGIAEGSAV AVGRYPEDVY QGGNPWYLAT
  351 AAAAEQLYDA IYQWKKIGSI SITDVSLPFF QDIYPSAAVG TYNSGSTTFN
  401 DIISAVQTYG DGYLSIVEKY TPSDGSLTEQ FSRTDGTPLS ASALTWSYAS
  451 LLTASARRQS VVPASWGESS ASSVPAVCSA TSATGPYSTA TNTVWPSSGS
  501 GSSTTTSSAP CTTPTSVAVT FDEIVSTSYG ETIYLAGSIP ELGNWSTASA
  551 IPLRADAYTN SNPLWYVTVN LPPGTSFEYK FFKNQTDGTI VWEDDPNRSY
  601 TVPAYCGQTT AILDDSWQ
```

Show predicted peptides also

Sort Peptides By ◉ Residue Number ○ Increasing Mass ○ Decreasing Mass

Start - End	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Sequence
73 - 83	1435.5845	1434.5772	1434.6204	-30	0	R.SDPNYFYSWTR.D (No match)
73 - 83	1435.5845	1434.5772	1434.6204	-30	0	R.SDPNYFYSWTR.D (Ions score 47)
91 - 101	1210.6034	1209.5961	1209.6394	-36	0	K.YLVDAFIAGNK.D (No match)
91 - 116	2927.4656	2926.4583	2926.5076	-17	1	K.YLVDAFIAGNKDLEQTIQQYISAQAK.V (No match)
102 - 116	1735.8329	1734.8256	1734.8788	-31	0	K.DLEQTIQQYISAQAK.V (No match)
102 - 116	1735.8329	1734.8256	1734.8788	-31	0	K.DLEQTIQQYISAQAK.V (Ions score 78)
117 - 136	1956.9371	1955.9298	1955.9800	-26	0	K.VQTISNPSGDLSTGGLGEPK.F (No match)
223 - 233	1114.5933	1113.5860	1113.6142	-25	0	R.ALVEGNALATR.L (No match)
274 - 302	3124.3113	3123.3040	3123.3663	-20	0	K.DVNSILGSIHTFDPAGGCDDSTFQPCSAR.A (No match)
274 - 302	3124.3113	3123.3040	3123.3663	-20	0	K.DVNSILGSIHTFDPAGGCDDSTFQPCSAR.A (Ions score 76)
316 - 334	1834.9109	1833.9036	1833.9585	-30	0	R.SIYAINSGIAEGSAVAVGR.Y (No match)
316 - 334	1834.9109	1833.9036	1833.9585	-30	0	R.SIYAINSGIAEGSAVAVGR.Y (Ions score 68)
420 - 433	1587.6808	1586.6735	1586.7213	-30	0	K.YTPSDGSLTEQFSR.T (No match)
420 - 433	1587.6808	1586.6735	1586.7213	-30	0	K.YTPSDGSLTEQFSR.T (Ions score 58)
434 - 457	2440.1653	2439.1580	2439.2281	-29	0	R.TDGTPLSASALTWSYASLLTASAR.R (No match)
434 - 457	2440.1653	2439.1580	2439.2281	-29	0	R.TDGTPLSASALTWSYASLLTASAR.R (Ions score 70)
555 - 580	2947.3777	2946.3704	2946.4076	-13	0	R.ADAYTNSNPLWYVTVNLPPGTSFEYK.F (No match)
584 - 598	1759.7350	1758.7277	1758.7809	-30	0	K.NQTDGTIVWEDDPNR.S (No match)




```
LOCUS       CAC28076                 618 aa            linear   PLN 14-NOV-2006
DEFINITION  glucoamylase [Talaromyces emersonii].
ACCESSION   CAC28076
VERSION     CAC28076.1  GI:12666724
DBSOURCE    embl accession AJ304803.1
KEYWORDS    .
SOURCE      Talaromyces emersonii (anamorph: Geosmithia emersonii)
  ORGANISM  Talaromyces emersonii
            Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina;
```

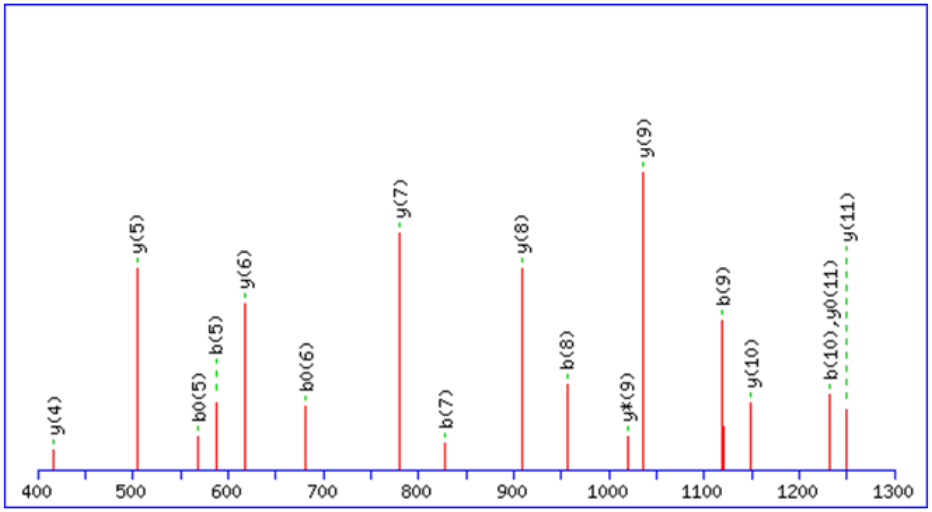
MATRIX SCIENCE Mascot Search Results

Peptide View

MS/MS Fragmentation of **DLEQTIQQYISAQAK**
Found in **gi|12666724**, glucoamylase [Talaromyces emersonii]

Match to Query 38: 1734.825624 from(1735.832900,1+) intensity(0.0000) index(3)
Title: Label: D3, Spot_Id: 202119, Peak_List_Id: 441379, MSMS Job_Run_Id: 21428, Comment:
Data file ppw_D3_130376210200.txt

Click mouse within plot area to zoom in by factor of two about that point
Or, Plot from 400 to 1300 Da Full range
Label all possible matches ○ Label matches used for scoring ◉



Monoisotopic mass of neutral peptide Mr(calc): 1734.8788
Fixed modifications: Carbamidomethyl (C) (apply to specified residues or termini only)
Ions Score: 78 Expect: 0.00024

#	Immon.	a	a*	a^0	b	b*	b^0	Seq.	y	y*	y^0	#
1	88.0393	88.0393		70.0287	116.0342		98.0237	D				15
2	86.0964	201.1234		183.1128	229.1183		211.1077	L	1620.8592	1603.8326	1602.8486	14
3	102.0550	330.1660		312.1554	358.1609		340.1503	E	1507.7751	1490.7486	1489.7645	13
4	101.0709	458.2245	441.1980	440.2140	486.2195	469.1929	468.2089	Q	1378.7325	1361.7060	1360.7219	12
5	74.0600	559.2722	542.2457	541.2617	587.2671	570.2406	569.2566	T	1250.6739	1233.6474	1232.6634	11
6	86.0964	672.3563	655.3297	654.3457	700.3512	683.3246	682.3406	I	1149.6263	1132.5997	1131.6157	10
7	101.0709	800.4149	783.3883	782.4043	828.4098	811.3832	810.3992	Q	1036.5422	1019.5156	1018.5316	9
8	101.0709	928.4734	911.4469	910.4629	956.4684	939.4418	938.4578	Q	908.4836	891.4571	890.4730	8
9	136.0757	1091.5368	1074.5102	1073.5262	1119.5317	1102.5051	1101.5211	Y	780.4250	763.3985	762.4145	7
10	86.0964	1204.6208	1187.5943	1186.6103	1232.6157	1215.5892	1214.6052	I	617.3617	600.3352	599.3511	6
11	60.0444	1291.6529	1274.6263	1273.6423	1319.6478	1302.6212	1301.6372	S	504.2776	487.2511	486.2671	5
12	44.0495	1362.6900	1345.6634	1344.6794	1390.6849	1373.6583	1372.6743	A	417.2456	400.2191		4
13	101.0709	1490.7486	1473.7220	1472.7380	1518.7435	1501.7169	1500.7329	Q	346.2085	329.1819		3
14	44.0495	1561.7857	1544.7591	1543.7751	1589.7806	1572.7540	1571.7700	A	218.1499	201.1234		2
15	101.1073							K	147.1128	130.0863		1

Seq	ya	yb	Seq	ya	yb	Seq	ya	yb
LE	215.1390	243.1339	LEQ	343.1976	371.1925	LEQT	444.2453	472.2402
LEQTI	557.3293	585.3243	LEQTIQ	685.3879	713.3828	EQ	230.1135	258.1084
EQT	331.1612	359.1561	EQTI	444.2453	472.2402	EQTIQ	572.3039	600.2988
QT	202.1186	230.1135	QTI	315.2027	343.1976	QTIQ	443.2613	471.2562
QTIQQ	571.3198	599.3148	TI	187.1441	215.1390	TIQ	315.2027	343.1976
TIQQ	443.2613	471.2562	TIQQY	606.3246	634.3195	IQ	214.1550	242.1499
IQQ	342.2136	370.2085	IQQY	505.2769	533.2718	IQQYI	618.3610	646.3559
QQ	229.1295	257.1244	QQY	392.1928	420.1878	QQYI	505.2769	533.2718
QQYIS	592.3089	620.3039	QQYISA	663.3461	691.3410	QY	264.1343	292.1292
QYI	377.2183	405.2132	QYIS	464.2504	492.2453	QYISA	535.2875	563.2824

MS/MS分数算法

- 不同于PMF打分原则，将匹配离子打分，再得到蛋白分数。

  算法复杂

检索结果分析

- 一个点鉴定得到多个蛋白的选择：一般选择得分最高的蛋白，但是如果得分最高的蛋白功能不明确（**unknown or hypothetical**），同时得分稍低一些的蛋白与最高分差别不大且功能相对明确，也可以选择得分稍低一些的蛋白。
- **MS/MS**以**peptide summary**结果为主，碎片离子可信度高，再结合**protein summary**分析。
- 看碎片离子峰匹配状况，匹配越好越可信
- 鉴定到的蛋白可以结合其它**(**如**pI, Mw)**等信息。有的鉴定到的结果差别较大：很多翻译后修饰、蛋白提取及电泳过程中的人为修饰以及蛋白降解、可变剪切等造成此现象，与质谱鉴定结果的好坏无关。

一级质谱鉴定和二级质谱鉴定如何选择

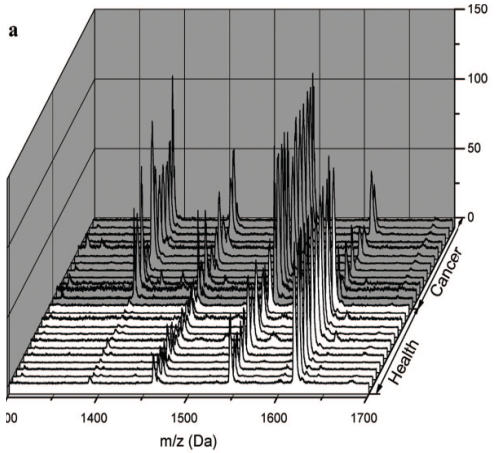
- 首先，一级质谱专一性信息少，只适用于有完整非冗余基因组序列或者蛋白序列的物种鉴定，对于大库检索（比如哺乳动物总库），随着库容量的增加，其特异性越差，结果越不可靠。
- 二级质谱提供的专一性信息多，即使对于大库检索，其鉴定结果也很可靠。二级质谱鉴定所有物种（包括有完整基因组信息的测序物种）也是一种质谱鉴定的趋势

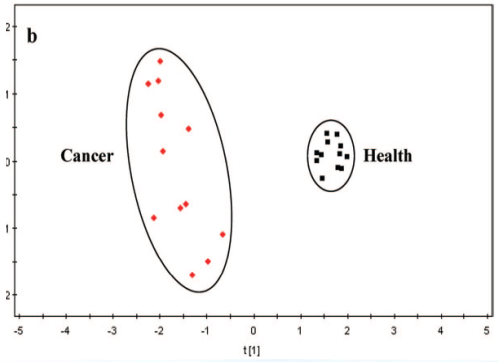
Denovo

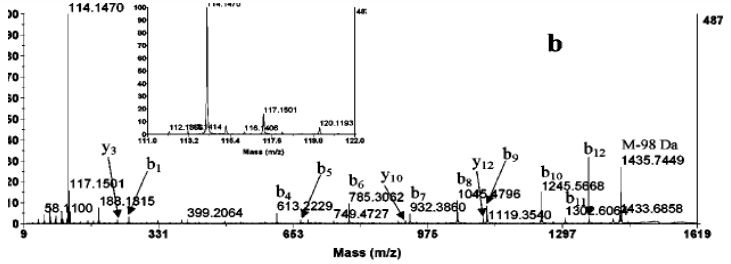
- 难，对样品、仪器要求高
- 解析MS/MS谱图要很有经验

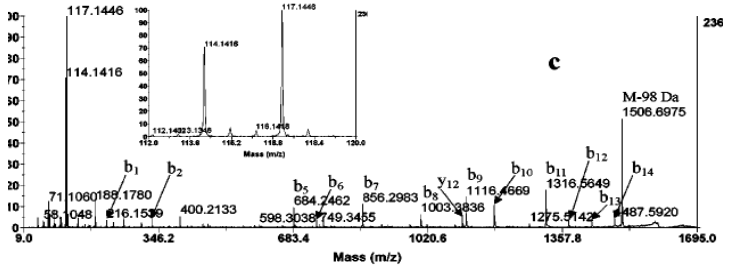
MALDI-TOF/TOF MS应用举例

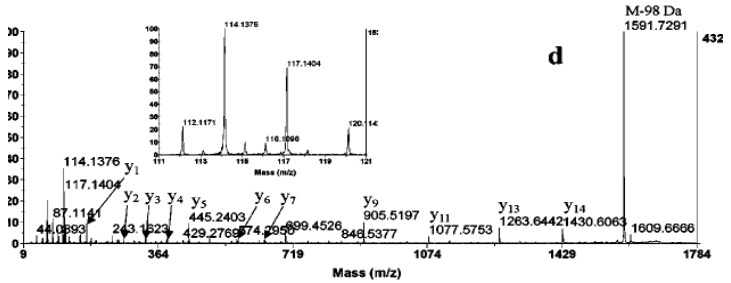
肝癌/健康人血清中磷酸化多肽表达的差异分析

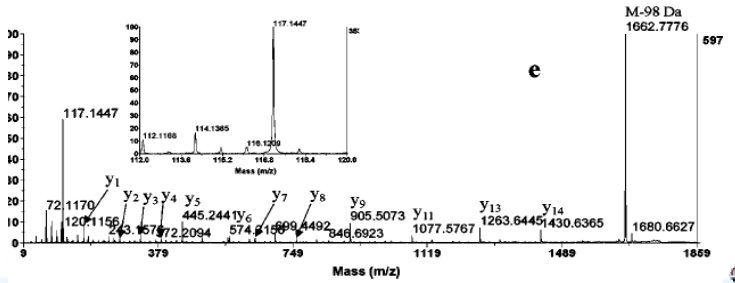




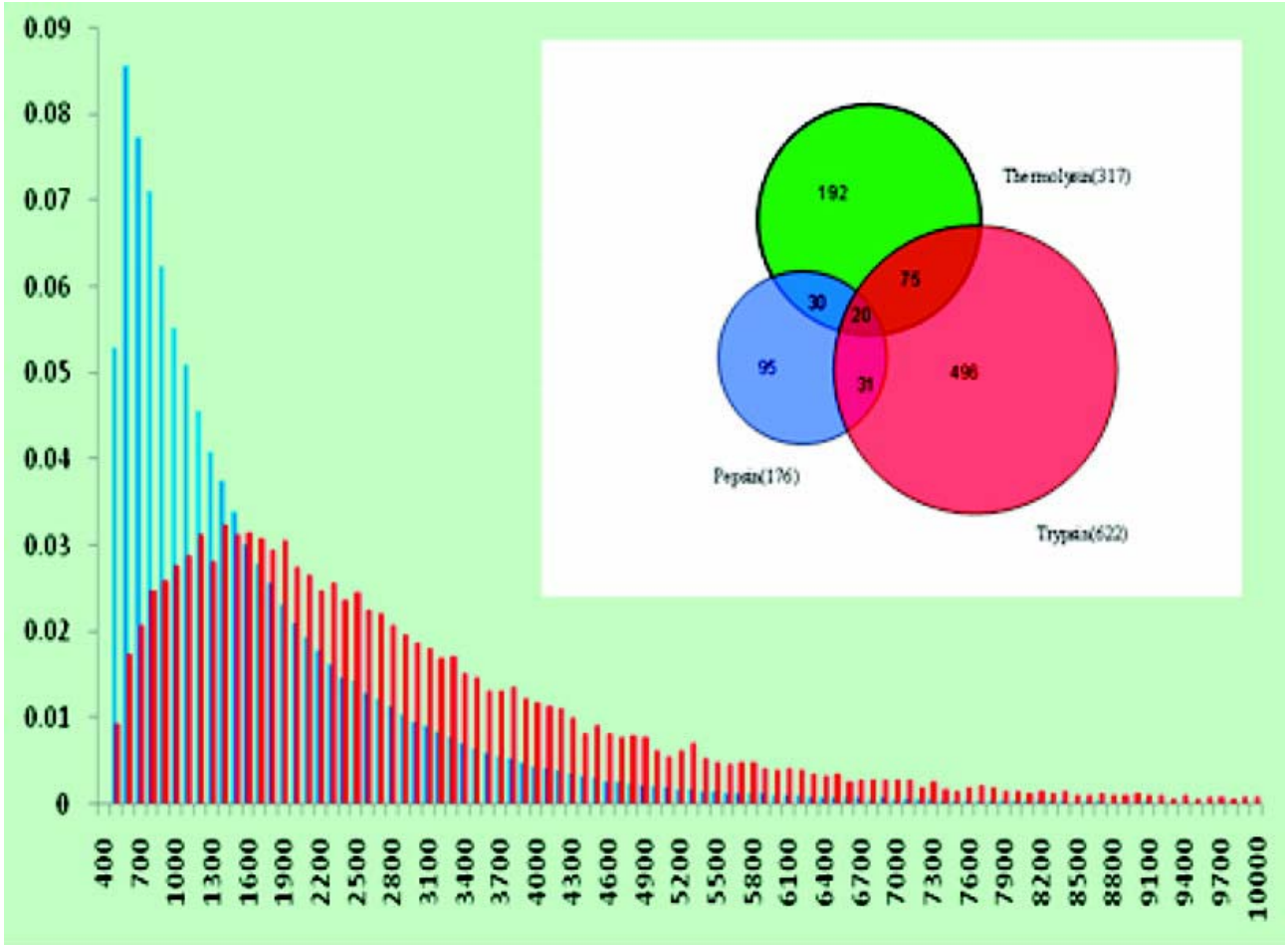








人肝蛋白糖基化蛋白和位点的鉴定



Distinguished groups

- **Ruedi Lab, ETH Zürich**
  the Institute for Systems Biology
- **Mann Lab, University of Southern Denmark**
  Max Planck Institute of Biochemistry
- **Gygi Lab, Harvard University**
- **Yates Lab, the Scripps Research Institute**
- **Heck Lab, University Utrecht**

质谱仪器使用选择

谢谢大家 ！

64806023

shutaosun@im.ac.cn