



华中农业大学
HUAZHONG AGRICULTURAL UNIVERSITY

亚蛋白质组学研究方法

报告人：吕芳

报告日期：2013-10-23

研究意义： 由于目前的技术条件无法对蛋白质进行完整的分析，较好的策略是分而治之，逐个击破。将蛋白质分成一个个亚蛋白质组（Subproteome），然后进行分析，由于样品的复杂度降低很多因此更可能发现有意义的低丰度蛋白。

翻译后修饰

(Post-translational modification, PTM)

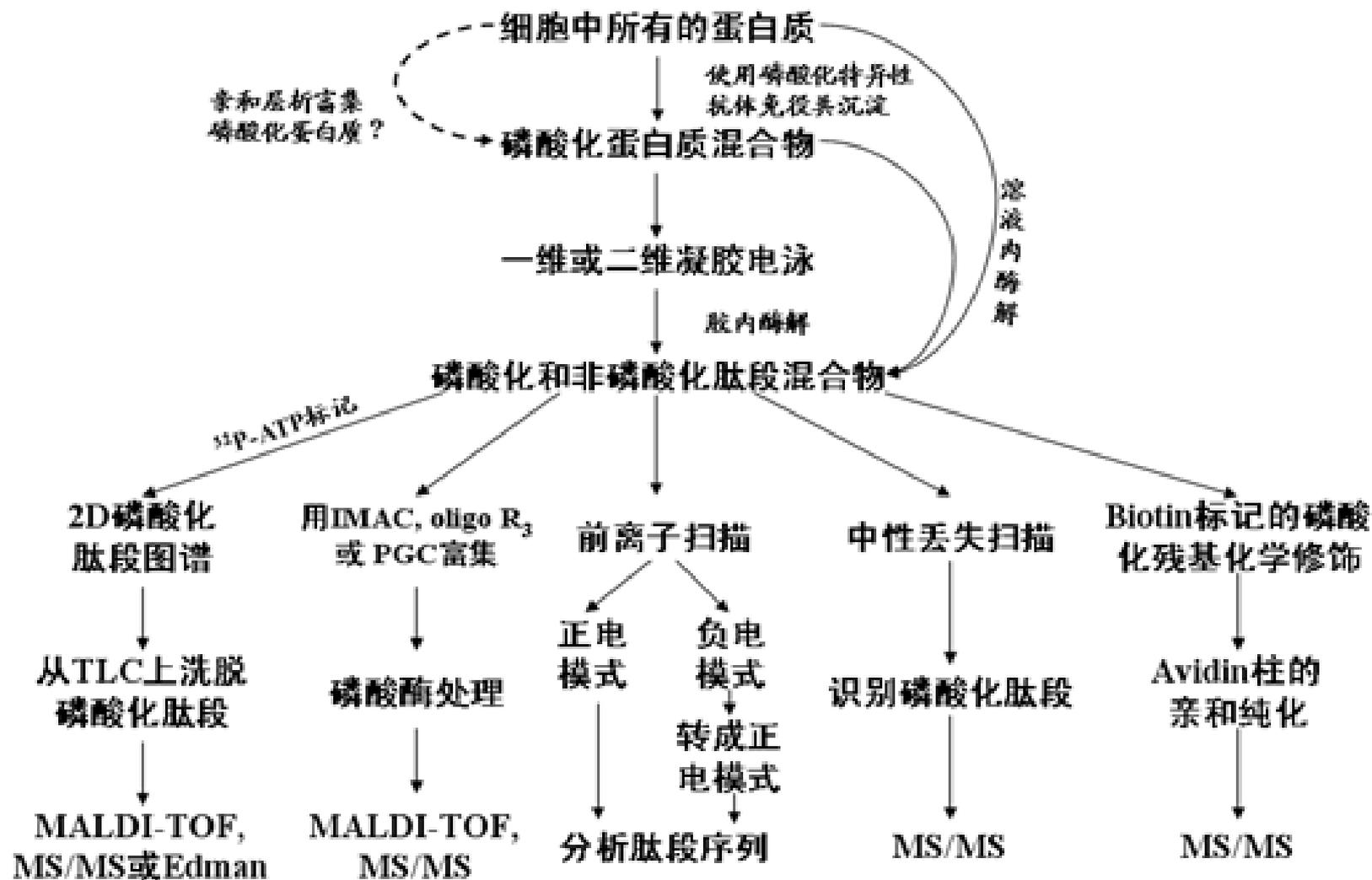
指蛋白质在翻译后的化学修饰，化学修饰是蛋白质加工的主要方式，修饰类型也很多，目前已知的翻译后修饰多达400余种。



研究方法

常规实验方法（糖基化，磷酸化）

生物信息学方法



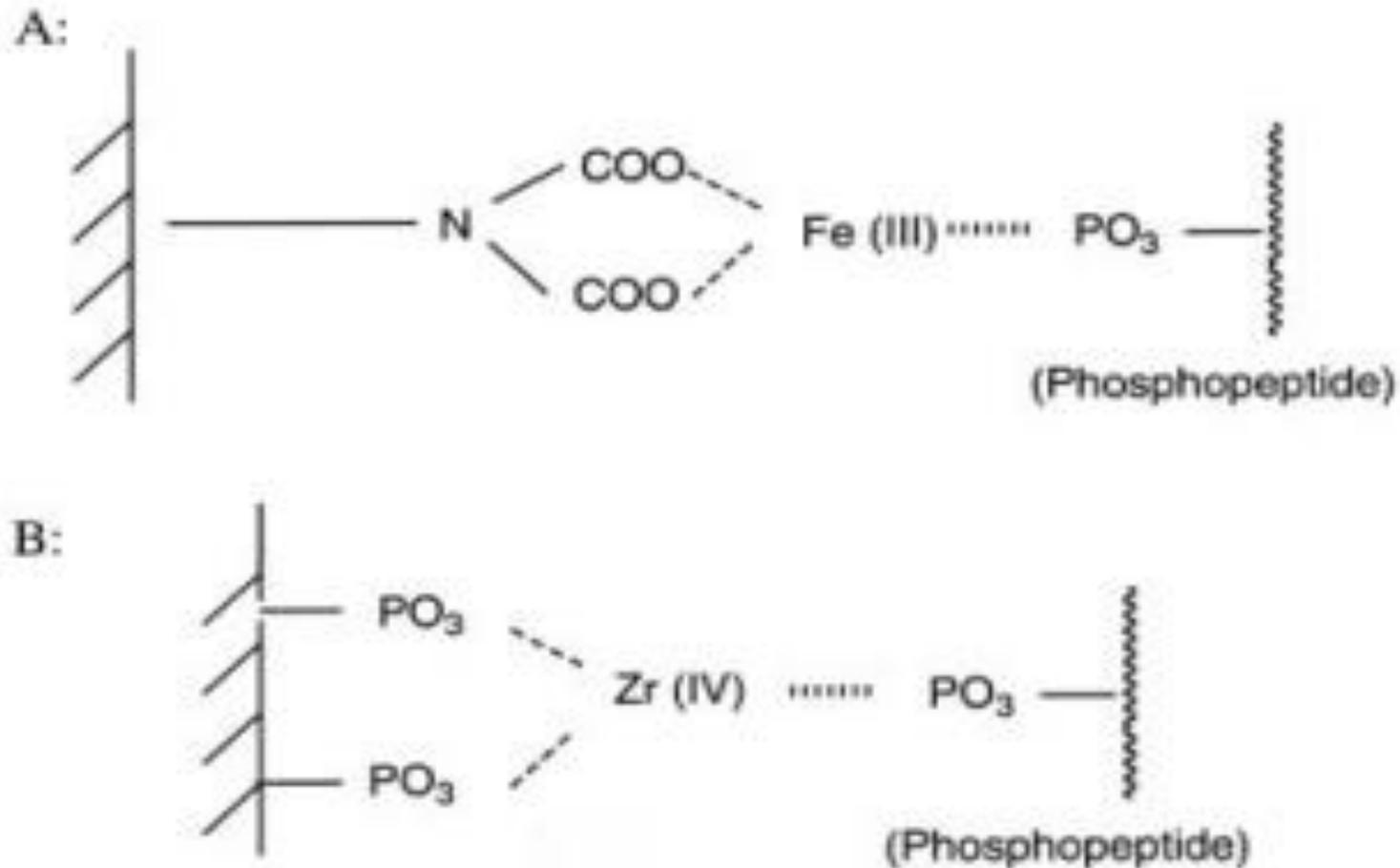


Figure 2. Schematic illustration of (A) conventional Fe^{3+} -IMAC and (B) novel Zr^{4+} -IMAC adsorbents for binding of phosphopeptides (Jiang, 2008)



MALDI-TOF-MS (Matrix-Assisted Laser Desorption/Ionization Time of Flight Mass Spectrometry) 是近年来发展起来的一种新型的软电离生物质谱，其无论是在理论上还是在设计上都是十分简单和高效的。

前离子扫描(precursor / parent ion scan)

利用磷酸基团的化学性质，检测它在各种条件下生成的有特点的离子，这其中有带负电荷的磷酸根离子，也有带正电荷的亚氨离子。该方法的离子源是电喷雾的，而质量分析器主要是串联质谱。

中性丢失扫描(neutral loss scan)

在三级四极杆串联质谱仪中，使第一个质量分析器 (Q1) 和第三个 (Q3) 质量分析器同步扫描，但扫描范围保持一个特定的电压差值，这个电压差所代表的 m/z 值为中性磷酸分子的质量(m/z 97.9或 m/z 49)。

蛋白质糖基化

是蛋白质翻译后的一种重要的加工过程，它是指蛋白质与葡萄糖（醛基糖）之间发生的非酶反应而生成糖化蛋白的过程，即指在肽链生物合成的同时或合成后在酶的催化下糖链被接到肽链上的特定糖基化位点的过程。

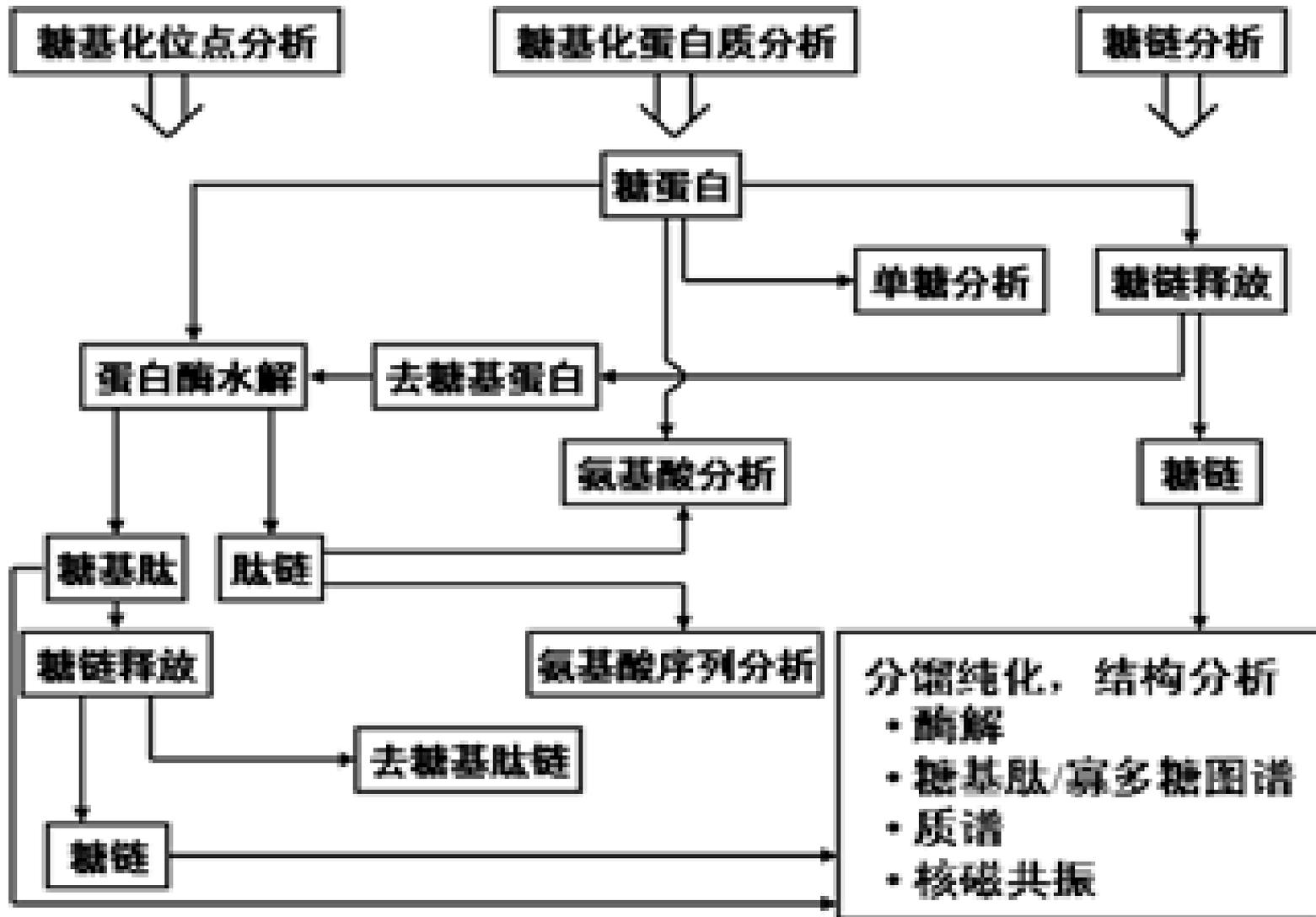


表 1 蛋白质翻译后修饰相关数据库

Table 1 protein post-translational modification databases

Database	Data	Statistics[13]
Swissprot ^[10]	Experiment verified ptms Potential ptms	11025 PTM sites on 4921 proteins 72308 PTM sites on 31026 proteins
Phosphoelm ^[11]	Verified phosphorylated sites in eukaryotes	5314 phosphorylated sites on 1805 proteins
Phosphosite ^[12]	Experiment verified ptms	
O-glycbase ^[14]	Verified glycosylated sites	242 glycosylated proteins
Resid ^[5]	Ptm types	424 ptm types
Dbptm ^[13]	Predicted pmts by HM M	14057 known PTM sites 772154 predicted PTM sites
Phosida ^[3]	Phosphorylated peptides from M S	6600 phosphorylated sites on 2244 proteins

表 2 翻译后修饰预测软件

Table 2 Prediction software of post-translational

Software	Prediction aim	
Netphos ^[17]	Phosphorylated sites on S, T, Y in eukaryotes	Feedforward
Predikin ^[18]	Substrates of ser/thr kinases	Similar
Scansite ^[19]	Potential phosphorylated sites by specific kinases	Position-specific

Netphosk ^[20]	Phosphorylated sites in eukaryotes by specific kinases	Feedforward neural networks
Disphos ^[21]	Phosphorylated sites by ser/thr/tyr kinases	Position-specific scoring matrix and disorder information
Predphospho ^[22]	Phosphorylated sites by four kinase groups	Support vector machine
Kinasephos ^[23]	Phosphorylated sites	Hidden markov model
Netoglyc ^[24]	O-glycosylation sites of mammals	Neural networks
Elm ^[25]	Consensus patterns in eukaryotes	Consensus pattern
Prosite ^[26]	Ptm consensus patterns, using motifscan	Consensus patterns
Netacet ^[27]	N-terminal acetylation sites	Neural networks
Kinasephos-like sulfation ^[13]	Sulfation sites	Hidden markov model
Sulfinator ^[28]	Tyrosine sulfation sites	Hidden markov model
Myristoylator ^[29]	N-myristoylation	Neural networks
Gpi-som ^[30]	Gpi-anchored signal	Self-organizing map

基于序列预测翻译后修饰的方法

- 保守模式（consensus patterns）
- 位点权重矩阵（position-specific scoring matrix PSSM）
- 隐马尔可夫链（profile hidden Markov）
- 神经网络（neuralnetworks）
- 支持向量机（supportvector machines SVM）

保守模式 (consensus patterns)

对发生同种修饰的序列片段进行多重比对，得到保守的特征序列片段。这些保守的氨基酸位点可能对发生修饰时蛋白的正确构象起关键作用可用来预测未知蛋白是否可能有相同的修饰

位点权重矩阵

(position-specific scoring matrix PSSM)

PSSM 是目前常用的从已有序列提取保守特征片段motif 并在目标序列上搜索motif的方法，矩阵的各元素表示相应位点出现相应氨基酸的概率，可直观的用序列标识图Sequence Logo表示。

基于质谱鉴定蛋白质翻译后修饰

- 数据库搜索法（Database Searching）
- 从头测序法（DeNovo Sequencing）

一、数据库搜索法

首先将数据库中的蛋白理论酶解，产生在一定误差范围内和母离子质量匹配的候选肽段，然后比较实验质谱和候选肽段的理论质谱，为实验质谱指派肽段，并用合适的打分函数给出分值。常用的数据库搜索法有SEQUEST、Mascot、X!Tandem等。



二、从头测序法

利用实验质谱的信息重建其对应的肽段序列，从而可能发现数据库中不存在的肽段。



质谱鉴定蛋白质翻译后修饰的质量控制

- PeptideProphet
- AScore



華中農業大學
HUAZHONG AGRICULTURAL UNIVERSITY

Thank you!