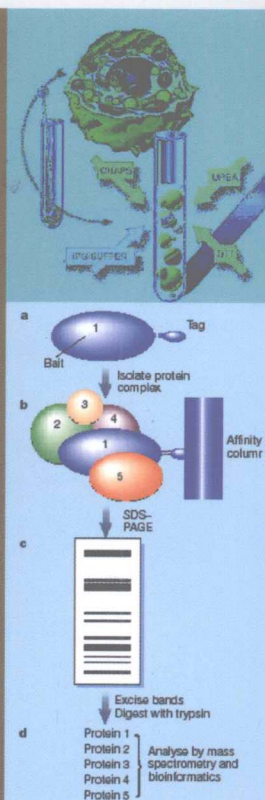
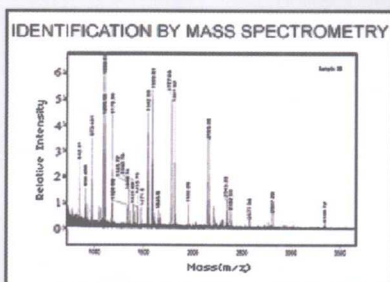
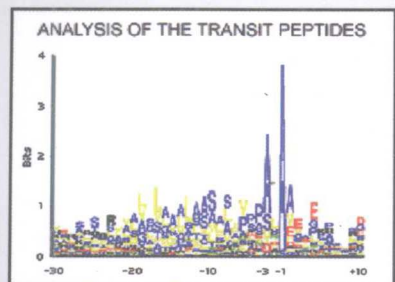
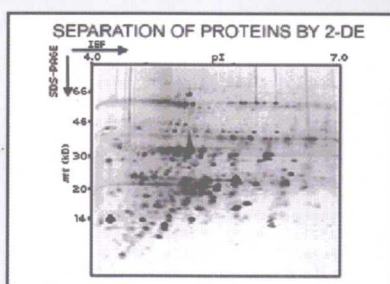
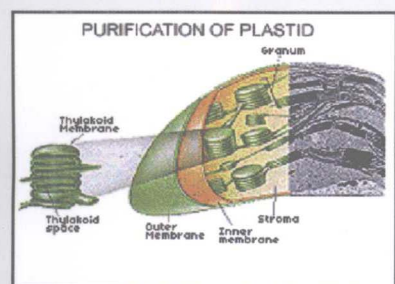


植物蛋白质组学

Zhiwu

Danbaizhi Zuxue

阮松林 马华升 主编



植物蛋白质组学

阮松林 马华升 主编

中国农业出版社

图书在版编目 (CIP) 数据

植物蛋白质组学/阮松林, 马华升主编. —北京: 中国农业出版社, 2009. 4

ISBN 978 - 7 - 109 - 13751 - 6

I. 植… II. ①阮…②马… III. 植物—蛋白质—基因组—研究 IV. Q946.1

中国版本图书馆 CIP 数据核字 (2009) 第 036226 号

中国农业出版社出版

(北京市朝阳区农展馆北路 2 号)

(邮政编码 100125)

责任编辑 徐建华

中国农业出版社印刷厂印刷 新华书店北京发行所发行

2009 年 4 月第 1 版 2009 年 4 月北京第 1 次印刷

开本: 787mm×1092mm 1/16 印张: 13.75

字数: 246 千字 印数: 1~1 000 册

定价: 50.00 元

(凡本版图书出现印刷、装订错误, 请向出版社发行部调换)

编写人员

主 编：阮松林 马华升

副主编：陶 梅 刘实忠

编写人员（按拼音排列）：

樊龙江（浙江大学）

刘实忠（海南热带作物学院）

马华升（杭州市农业科学研究院）

阮松林（杭州市农业科学研究院）

孙玉强（浙江大学）

陶 梅（中国农业科学院）

王 杰（国家生物医学分析中心）

祝水金（浙江大学）

邹俊杰（中国农业大学）

主要作者介绍

.....
□□□□□□□□□□□□□□□□□□□□

阮松林，男，浙江省上虞市人。2001 年 7 月在浙江大学获得农学博士学位。2002—2005 年在中国农业大学生物学院植物生理学与生物化学国家重点实验室从事博士后研究。2004 年获得国家自然科学基金资助。现主持浙江省自然科学基金、杭州市重大科技创新等项目。2006—2007 年分别列入杭州市“131”和浙江省“151”新世纪人才工程培养计划。2007 年 4 月获得“第八届杭州市青年英才”称号及五四青年奖章。现为杭州市农科院生物技术研究所副所长、副研究员、硕士生导师，植物分子生物学与蛋白质组学实验室（PMBP Lab）负责人。主要从事植物逆境蛋白质组学研究，已申请国家专利 7 项（其中申请基因专利 6 项，授权 1 项），发表论文 30 余篇（其中 SCI 收录论文 6 篇），参编著作 3 部。

主要作者介绍

.....
oooooooooooooooooooo

马华升，男，浙江省东阳市人，享受国务院特殊津贴。杭州市“131”新世纪人才工程第一层次和浙江省“151”跨世纪人才工程第二层次培养对象。现为杭州市农科院生物技术研究
所所长、研究员、硕士生导师。曾从事杂交水稻制繁种技术研究
和新品种（组合）引选、试验示范推广工作，现主要从事植物
逆境蛋白质组学及分子生物学研究。先后主持、参加国家、
省、市级科研项目 20 余项，获各级科技进步奖 30 余项，其中
获省部级以上 8 项。申请国家专利 8 项（其中申请基因专利 6
项），已授权 1 项。发表论文 40 余篇（其中 SCI 收录论文 3
篇），主、参编著作各 1 部。

序

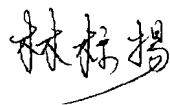
蛋白质组学是在基因组学研究取得重大成就和高通量蛋白质分析技术实现突破的背景下产生的一门新兴学科。基因组研究的快速发展特别是基因组全测序的完成是蛋白质组学研究的重要基础。植物蛋白质组学的研究主要进展集中在模式植物拟南芥 (*Arabidopsis thaliana*) 上。拟南芥 (*Arabidopsis thaliana*) 的基因组序列的完成奠定了它蛋白质组学研究的基础。多国拟南芥研究指导委员会 (Multinational *Arabidopsis* Steering Committee, MASC) 在 2006 年建立了拟南芥蛋白质组学的亚委员会 (MASCP), 并建立了整合的数据库和技术标准 (www.masc-proteomics.org)。水稻 (*Oryza sativa*) 和玉米 (*Zea mays*) 两种重要经济作物的蛋白质组学研究也方兴未艾 (ppdb.tc.cornell.edu; cdna01.dna.affrc.go.jp/RPD)。植物蛋白质组学的研究是植物系统生物学的必不可少的一部分, 是理解植物生物学的关键, 它将为解决和改善人类营养和环境, 地球生态等做出重大的贡献。

生物功能主要执行者是蛋白质, 而蛋白质有其自身特有的活动规律包括蛋白质修饰加工、转运定位、结构变化、蛋白质与蛋白质间、蛋白质与其他生物大分子的相互作用等。因此蛋白质组学的研究包括了蛋白定量, 蛋白翻译后修饰 (磷酸化, 糖基化等), 蛋白-蛋白相互作用等的研究。这部书稿较为全面系统地介绍了这些蛋白质组学研究所需的关键技术及其在植物生物学研究中的应用。书稿总体布局合理, 所涉及内容丰富且具有新颖性, 表述清晰, 能全面反映当前的蛋白质组学研究动态, 使读者阅后既能了解蛋白质组学前沿技术又能及时把

握植物蛋白质组学研究最新动态。

全书共十四章，主要从三个层面展开。在书后第一层面着重介绍蛋白质组概念、产生背景、发展趋势和重要性，勾画出蛋白质组学从产生到发展的清晰轮廓。第二层面主要介绍蛋白质组学关键技术，包括双向电泳、高效液相色谱、蛋白芯片、生物质谱、定量蛋白质组学技术、蛋白质与蛋白质相互作用和蛋白质组学相关数据库，为读者展现当今蛋白质组学研究的最新技术。第三层面综述了蛋白质组学技术在植物生物学研究中的应用及研究进展，主要从植物遗传多样性、植物发育、植物逆境和蛋白标记与 PQL 等方面作详细概括性的介绍，让读者准确把握植物蛋白质组学最新研究动态。著作在介绍前人研究结果的同时，在一些章节中编入了作者最新的研究成果，以着力体现本书的新颖性和实用性。书中最后章节详细介绍了双向电泳和质谱技术操作步骤，书后附加蛋白质组学常用数据库说明及网址，便于初学者学习参考。

显然，在当今蛋白质组学技术不断改进与成功应用研究中，本书必将成为广大从事植物蛋白质组学研究的科研人员了解和掌握蛋白质组学技术的实用参考工具，对今后的科研工作具有重要的指导作用。



浙江大学浙江加州国际纳米研究院系统生物学

首席科学家，教授，美国华盛顿大学教授

前 言

2000年6月，人类基因组计划完成，使得基因组学研究从结构转到功能分析，预示着后基因组学即功能基因组学时代的到来。同时，人们清楚地意识到通过基因组结构分析无法全面了解生命活动实现者——蛋白质的活动规律，如蛋白质修饰加工、转运定位、结构变化、蛋白质与蛋白质间、蛋白质与其他生物大分子的相互作用等，均无法在基因组水平上获得。因此，需要建立一套高通量蛋白质分析技术对细胞内蛋白质表达和功能模式进行直接分析。生物质谱技术的快速发展为生物大分子蛋白质的高通量鉴定成为可能，而且双向电泳、液相色谱、定量蛋白质技术改进与建立为蛋白质实现高通量及定量分析提供了技术保障。在此背景下，一门新兴学科——蛋白质组学应运而生。

随着模式植物拟南芥和水稻测序计划相继完成，植物蛋白质组学研究真正进入了一个快速发展时期，每年有成百上千篇关于植物蛋白质组学的原始论文发表在相关杂志上，研究内容涉及植物生物学重点领域，包括植物发育、非生物逆境、营养、植物与微生物或昆虫互作等方面。而且通过蛋白质组学的系统研究，发现了许多具有重要生物学功能的新蛋白质，有助于进一步揭示植物生命活动规律。随着植物新蛋白质的发现及功能解析，使得植物蛋白质数据库不断被补充完善，而且还建立植物蛋白质新数据库，丰富了蛋白质数据库资源，对植物蛋白质鉴定、近缘种蛋白质组分析及新育种目标的确定具有十分重要的意义。

为了将蛋白质组关键技术和植物蛋白质组学研究进展系统地介绍

给广大大专院校师生和从事相关领域的科研工作者，我们以 2006 年 11 月和 12 月的遗传杂志上发表的两篇综述文章为基础，不断充实相关内容，使之具有系统性和新颖性，让广大读者既能了解蛋白质组前沿技术又能及时把握植物蛋白质组学研究最新动态，希望《植物蛋白质组学》一书真正成为您的科研工作重要参考书。全书共十四章，其中第一至第七章主要介绍蛋白质组概念、产生背景、重要性和蛋白质组学关键技术（包括双向电泳、高效液相色谱、蛋白芯片、生物质谱、定量蛋白质组学技术、蛋白质与蛋白质相互作用和蛋白质组学相关数据库）；第八至第十三章主要介绍蛋白质组学技术在植物生物学研究中的应用及研究进展，研究内容涉及植物遗传多样性、植物发育（包括种子发育）、植物逆境（包括生物逆境和非生物逆境）和蛋白标记与 PQL；第十四章主要介绍双向电泳和质谱技术操作指南及相关说明，便于初学者了解双向电泳技术和质谱技术的一般操作程序。另外，本书增加了附录，包括双向电泳完整的操作步骤、聚丙烯酰胺凝胶电泳凝胶的配置、细胞样品的一般处理步骤、组织样品的一般处理步骤及蛋白质和蛋白质组相关数据库，主要供负责具体操作的实验人员参考。本书适用于大专院校农林及生物专业的师生和相关科研院所的科技人员。

本书由编写组集体撰写完成，具体章节分工为：第一章，阮松林、马华升；第二章，刘实忠、阮松林；第三章，阮松林、王杰、马华升；第四章，阮松林、陶梅；第五章，阮松林；第六章，樊龙江、马华升、阮松林；第七章，马华升、祝水金、孙玉强；第八章，祝水金、孙玉强、马华升；第九章，祝水金、马华升、阮松林；第十章，陶梅、马华升、阮松林；第十一章，阮松林、马华升、邹俊杰；第十二章，马华升、刘实忠、阮松林；第十三章，孙玉强、马华升、阮松林；第十四章，邹俊杰、王杰。全书由阮松林和马华升统稿和定稿。

本书的出版受到国家自然科学基金项目（30300218）、浙江省自然科学基金项目（Y307016）、杭州市重大科技创新项目（20072312A03）和杭州市科研院所专项（20061922N04）等项目的资助，在此深表感谢。在撰写本书过程中，参考和引用了国内外的一些专著和论文资料和图表，在此向这些作者表示感谢。

美国华盛顿大学林标扬教授对本书内容修改提出许多建设性的意见，并在百忙中抽出时间欣然为本书作序，在此谨向他深表感谢。

最后，由于水平有限，在本书编写过程中，难免出现错误和不足之处，恳请同行专家和读者批评指正，不胜感激。

阮松林

二〇〇八年十二月于杭州

目 录

序

前言

第一章 蛋白质组学概论	1
第一节 蛋白质组概念	1
第二节 蛋白质组学产生背景	1
第三节 蛋白质组学研究内容	3
第四节 蛋白质组学的研究方法	4
第五节 蛋白质组学研究技术	4
第六节 蛋白质组学研究的最终目标	4
第七节 蛋白质组学的研究意义及应用	5
第八节 蛋白质组学发展趋势	6
参考文献	7
第二章 蛋白质组分离技术	10
第一节 双向电泳	10
第二节 高效液相色谱	18
第三节 蛋白质芯片	21
参考文献	25
第三章 蛋白质组鉴定技术	28
第一节 生物质谱技术	28
第二节 生物质谱的应用	33
第三节 展望	37
参考文献	37
第四章 定量蛋白质组学分析	41
第一节 荧光双向差显凝胶电泳 (DIGE) 技术	41

2	植物蛋白质组学	
第二节	同位素亲和标签技术	43
第三节	同位素标记相对和绝对定量 (iTRAQ) 技术	48
	参考文献	51
第五章	蛋白质与蛋白质的相互作用	53
第一节	酵母双杂交技术	53
第二节	串联亲和纯化技术	55
	参考文献	58
第六章	蛋白质组学相关的生物信息学	60
第一节	2-D 数据库	60
第二节	支持质谱数据的蛋白质鉴定的数据库	61
第三节	分析蛋白与蛋白相互作用的数据库	62
第四节	基于质谱技术的计算蛋白质组学	63
	参考文献	69
第七章	不同水平上的蛋白质组学	71
第一节	组织、器官水平上蛋白质组学	71
第二节	亚细胞水平上蛋白质组学	72
	参考文献	85
第八章	植物遗传多样性分析	89
第一节	种间的遗传变异性	89
第二节	种内的遗传变异性	90
第三节	品种鉴定	90
	参考文献	91
第九章	植物发育相关蛋白质组学	93
第一节	与发育性状相关突变体差异蛋白质组分析	93
第二节	与种子发育过程相关的蛋白质组分析	94
第三节	外源植物激素处理对植物发育过程蛋白质表达的影响	96
	参考文献	97
第十章	种子蛋白质组学研究	99
第一节	种子蛋白质及其分类	99

第二节 种子蛋白质与蛋白质组的研究内容	100
第三节 展望	105
参考文献	106
第十一章 植物非生物逆境相关蛋白质组学	111
第一节 植物热休克蛋白质组的鉴定与分析	111
第二节 植物冷响应蛋白质组的鉴定与分析	111
第三节 植物耐旱相关蛋白质组的鉴定与分析	114
第四节 植物耐盐相关蛋白质组的鉴定与分析	115
第五节 植物营养逆境响应的蛋白质组的鉴定与分析	127
第六节 植物对其他逆境响应的蛋白质组的鉴定与分析	128
参考文献	129
第十二章 植物生物逆境相关蛋白质组鉴定与分析	134
第一节 植物与病原菌互作的蛋白质组鉴定	134
第二节 植物抗虫蛋白质组的鉴定与分析	136
第三节 豆科植物与根瘤菌互作的蛋白质组鉴定与分析	136
参考文献	138
第十三章 蛋白标记遗传作图与候选蛋白	141
第一节 蛋白标记遗传作图	141
第二节 蛋白数量位点 (PQL) 和候选蛋白	143
参考文献	144
第十四章 蛋白质组学相关技术实验操作指南	147
第一节 双向电泳实验操作技术	147
第二节 蛋白质质谱分析操作技术	176
参考文献	179
附录 1 双向电泳完整的操作步骤	180
附录 2 聚丙烯酰胺凝胶电泳凝胶的配置	183
附录 3 细胞样品的一般处理步骤	185
附录 4 组织样品的一般处理步骤	186
附录 5 蛋白质和蛋白质组相关数据库	187
英文缩略词表	196

第一章 蛋白质组学概论

第一节 蛋白质组概念

蛋白质组 (Proteome) 一词是由英文单词蛋白质 (Protein) 的前半部分加上基因组 (Genome) 的后半部分组合而成, 意指 “一种基因组所表达的全套蛋白质”, 即包括一种细胞乃至一种生物所表达的全部蛋白质。蛋白质组概念最早是由澳大利亚学者马克·威尔金斯 (Marc Wilkins) 和凯斯·威廉姆斯 (Keith Williams) 于 1994 年意大利召开的一次科学会议上首次提出, 并由 Wasinger 等人于 1995 年首次在出版物中使用^[1]。它是指一个有机体的全部蛋白质组成及其活动方式。目前蛋白质组是指一个生物体基因组、一个细胞或组织所表达的全部蛋白质成分。与基因组不同的是, 蛋白质组作为一个整体, 在不同的时空条件下, 在一个生物体的不同组织中是不同的。而一个生物体仅有一个特定的基因组。一个生物体的蛋白组未必与基因组存在一一对应关系, 主要是由于基因存在转录后的不同剪接方式和翻译后蛋白的修饰所致。从基因表达的角度来看, 蛋白质组的蛋白质数目总是多于基因组的基因数目。从蛋白质修饰的角度来看, 蛋白质组的蛋白质数却多于其相应的开放阅读框 (open reading frame, ORF) 数目, 因为 mRNA 的剪切和编辑可使一个 ORF 产生数种蛋白质, 蛋白质翻译后的修饰, 如糖基化、磷酸化同样增加蛋白质的种类, 氨基酸序列一致的一级结构在一定条件下可以形成功能完全不一样的具有不同空间结构的蛋白质, 如朊病毒。因此, 蛋白质组研究是为了识别及鉴定一个细胞或组织所表达的全部蛋白质以及它们的表达模式 (钱小红, 1998)^[2]。蛋白质组学是以蛋白质组为研究对象, 从整体蛋白质水平上, 在一个更加深入、更加贴近生命本质的层次上去探索和发现生命活动的规律和重要的生理、病理现象等。

第二节 蛋白质组学产生背景

植物蛋白质组学是在基因组学的研究成就和高通量的蛋白质分析技术得到突破的背景下产生的一门新兴学科, 基因组研究的发展是蛋白质组学产生的重要前提。

基因组研究是生物科学近十几年来来的研究热点。人类基因组计划被誉为 20 世纪的三大科技工程之一, 并取得了辉煌的成就。2000 年 6 月科学家公布人类基因组工作草图 (Macil Wain 等, 2000 年)^[3], 标志着人类在解读自身 “生命之

书”的路上迈出了重要一步。2001年2月,中、美、日、德、法、英等6国科学家和美国塞莱拉公司联合公布人类基因组图谱及初步分析结果(Lander等,2001等)^[4],宣告了一个新的纪元——后基因组(即功能基因组)时代的到来。在后基因组时代,研究的重点已从揭示遗传信息转移到功能基因组学上来。但是,由于生物功能主要体现者是蛋白质,而蛋白质有其自身特有的活动规律。如蛋白质修饰加工、转运定位、结构变化、蛋白质与蛋白质间、蛋白质与其他生物大分子的相互作用等,均无法在基因组水平上获得。因为基因组学有这样的局限性,促使人们从整体水平上探讨细胞蛋白质的组成及其活动规律。

植物基因组学的研究主要集中在拟南芥(*Arabidopsis thaliana*)和水稻(*Oryza sativa*)两种模式植物上。2000年12月,美、英等国科学家宣布测出拟南芥基因组的完整序列(The Arabidopsis Genome Initiative,2000年),这是人类首次全部破译高等植物的基因序列。2002年是水稻基因组学研究取得重大成就的一年,首先中国科学家和Syngenta公司的科学家分别发表籼稻和粳稻基因组“工作框架图”(Yu et al.,2002年;Goff et al.,2002年)^[5,6],继后日本和中国的科学家又分别公布了粳稻第1号和第4号染色体的全序列(Feng et al.2002;Sasaki et al.,2002年)^[7,8]以及籼稻粳稻基因的“精细结构图”,被认为是基因组学研究的又一个重要里程碑。

基因组密码的破译,拉开了生命科学研究的序幕,但是,要真正揭示生命活动的奥秘,基因组研究本身又无能为力。因为,基因组仅仅是遗传密码和遗传信息的载体,在生命活动的不同过程中恒定不变,不能反映有机体在生命活动过程中基因表达的时空关系和网络调控。在后基因组时代,研究重心转移到基因功能的解析,即利用结构基因组所提供的信息和高通量的实验手段在转录组(Transcriptome)和蛋白质组水平上系统地分析基因的功能(Oliver,2000年)^[9]。

拟南芥和水稻的功能基因组学研究已经开始,其中美国和日本科学家做了大量的工作。其他植物如玉米(Touzot等,1996年)^[10]、小麦(Weiss等,1997年)^[11]、苜蓿(Catoira et al.,2000年)^[12]、松树(Costa等,1999年)^[13]等的功能基因组学研究也已经有人涉及。

蛋白质是基因功能的体现者和执行者。现在已经证明,一个基因并不只产生一个相应的蛋白质,它可能会产生几个,甚至几十个蛋白质。机体所处的不同环境和本身的生理状态差异,会导致基因转录产物有不同的剪切和转译成不同的蛋白。蛋白再进行加工修饰和转移定位,才具有活性和生物功能,产生相应的生理作用,适宜相应的生存环境(Li & Assmann,2000年)^[14]。在转录水平上所获取的基因表达的信息并不足以揭示该基因在细胞内的确切功能。直接对蛋白质的表达模式和功能模式进行研究就成为生命科学发展的必然趋势。因此,研究基因组编码的全蛋白质功能及其相互作用关系的蛋白质组学应运而生(Anderson和

Anderson, 1998 年)^[15]。尽管蛋白质组学在 20 世纪 90 年代中后期才出现,但由于学科的前沿性和巨大的应用市场,《Nature》、《Science》杂志在公布人类基因组序列草图的同时,分别发表了述评和展望,将蛋白质组学的地位提到前所未有的高度,认为它是功能基因组学前沿研究的战略制高点和新世纪最大的战略资源——“有用基因”争夺战的重要“战场”(Kaiser, 2000 年; Macil Wain, 2000 年)^[3,16]。

第三节 蛋白质组学研究内容

蛋白质组学研究主要包括蛋白质的表达模式和蛋白质的功能模式两个方面。

蛋白质表达模式的研究是蛋白质组学研究的基础内容,主要是研究特定条件下某一细胞或组织的所有蛋白质的表征问题。常规的方法是提取蛋白质,经 2-DE (Two-dimensional electrophoresis) 分离形成一个蛋白质组的二维图谱,通过计算机图像分析得到各蛋白质的等电点、分子量、表达量等,再结合以质谱分析为主要手段的蛋白质鉴定,建立起细胞或组织或机体在所谓“正常生理条件下”的蛋白质组图谱和数据库。然后,在此基础上,可以比较分析在变化了的条件下蛋白质组所发生的变化,如蛋白质表达量的变化、翻译后的加工修饰、蛋白质在亚细胞水平上定位的改变等,从而发现和鉴定出特定功能的蛋白质及其基因。对蛋白质组成的分析鉴定是蛋白质组学中 with 基因组学相对应的主要部分,它要求对蛋白质进行表征即实现所有蛋白质的分离、鉴定——亦即其图谱化。

蛋白质功能模式的研究是蛋白质组学研究的重要目标。无论是基因组研究还是蛋白质组研究,最终目标是揭示所有基因或蛋白质功能及其作用模式。一方面,蛋白质与蛋白质、蛋白质与 DNA 之间的相互作用、相互协调是细胞进行信号传导及代谢活动的基础。另一方面,蛋白质的结构是蛋白质发挥其功能的前提,对蛋白质结构的认识也成为了解大量涌现的新基因功能的一个重要途径。蛋白质间的相互作用归纳起来有以下几类: 1) 分子和亚基的聚合; 2) 分子杂交; 3) 分子识别; 4) 分子自组装; 5) 多酶复合体。通过分析一个蛋白质是否能和功能已知的蛋白质相互作用可得到揭示其功能的线索。

Humphery-Smith 等 (1997)^[17]总结了基因组结果后提出了“功能蛋白质组 (Functional Proteome)”的新概念。即细胞内与某个功能有关或在某种条件下的一群蛋白质。在此基础上,我国学者李伯良 (1998)^[18]提出了“功能蛋白质组学 (Functional Proteomics)”的概念。即把“功能蛋白质组”作为主要研究内容,从局部着手,注重研究那些可能涉及到特定功能机理的蛋白质全体。这一概念的提出,为蛋白质组研究的可能性奠定了理论基础。因此,从一个生物体的具有特定功能的器官、组织、细胞着手,在不同生理条件下研究表达蛋白质的功能。