

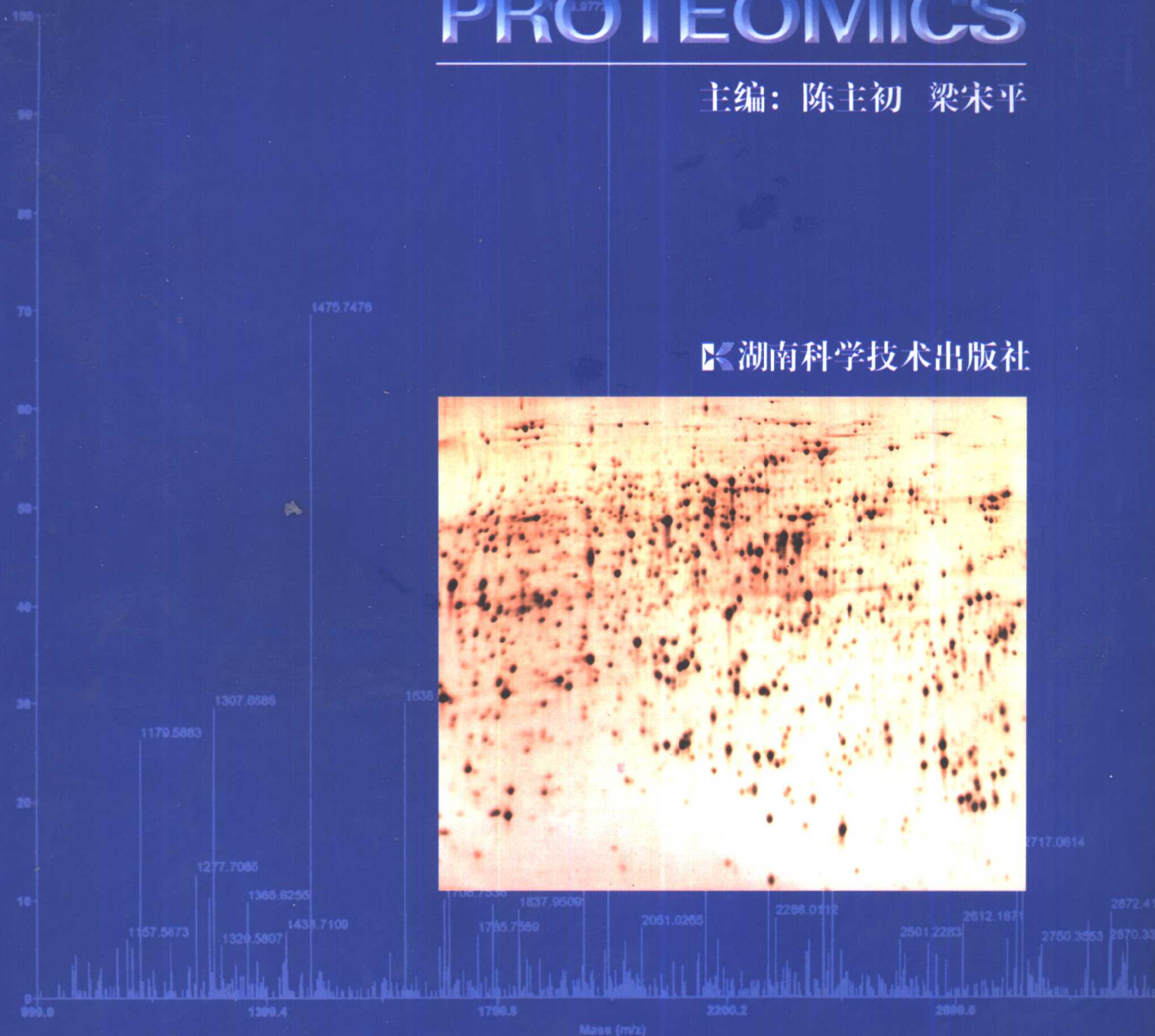
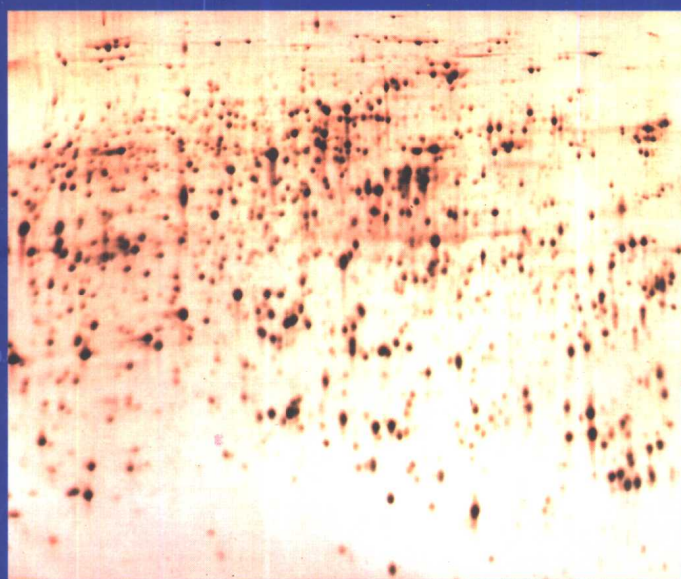
CANCER PROTEOMICS

肿瘤蛋白质组学

CANCER
PROTEOMICS

主编：陈主初 梁宋平

湖南科学技术出版社



CANCER PROTEOMICS

肿瘤蛋白质组学

**CANCER
PROTEOMICS**

主 编：陈主初 梁宋平

副主编：肖志强

编 者（按姓氏笔画排序）

王贤纯	关勇军	陈主初	陈 平
肖志强	李 峰	李 萃	张广森
梁宋平	谢锦云		

❧ 湖南科学技术出版社

肿瘤蛋白质组学

主 编：陈主初 梁宋平

责任编辑：梅志洁

出版发行：湖南科学技术出版社

社 址：长沙市湘雅路 280 号

<http://www.hnstp.com>

邮购联系：本社直销科 0731-4375808

印 刷：国防科技大学印刷厂

(印装质量问题请直接与本厂联系)

厂 址：长沙市砚瓦池正街 47 号

邮 编：410073

经 销：湖南省新华书店

出版日期：2002 年 9 月第 1 版第 1 次

开 本：787mm×1092mm 1/16

印 张：12.5

字 数：320000

书 号：ISBN 7-5357-3539-8/R·788

定 价：22.00 元

(版权所有·翻印必究)

序

生命科学的一个崭新时代——蛋白质组时代已经开始，蛋白质组学是继基因组研究之后的又一核心和具有广泛应用价值的学科。

众所周知，蛋白质组学是以蛋白质为研究对象，在蛋白质多肽谱和基因产物图谱技术的基础上发展起来的一门科学，是通过对基因表达产物——蛋白质进行动态和整体水平的研究来阐述环境、疾病、药物等对细胞代谢的影响，并分析其主要作用机制、解释基因表达调节的三要素方式。因此，它不是一种单纯的方法学，而是一个以蛋白质为核心揭示生命规律的研究新领域。可以相信，随着基因组研究的进一步拓展，蛋白质研究数据的不断积累，新方法、新技术的突破和生物信息学工具的完善，蛋白质组的研究一定能从幼稚走向成熟，为生命科学的研究、生物技术的应用和人类疾病的防治带来新的革命。此外，蛋白质组研究既有理论意义，又有潜在的应用价值，通过蛋白质组研究发现的有关蛋白质最有可能转化为应用，发挥社会效益和经济效益。

中南大学湘雅医学院陈主初教授和湖南师范大学梁宋平教授立足生命科学的最新前沿领域，组织从事该领域研究的中青年专家学者，在广泛收集国内外资料的基础上，结合各自的实践，在国内率先编写了《肿瘤蛋白质组学》一书。该书着重介绍了蛋白质组学的产生、基本原理和技术、肿瘤蛋白质组学的现状、研究策略与发展趋势。该书内容全面，资料新颖，以蛋白质组为主线，由浅入深，循序渐进，理论联系实际，是目前我国生命科学领域中第一本系统介绍蛋白质组研究技术及其在肿瘤研究中应用的书籍，具有较高的学术水平和实用价值，既可作为生命科学与医药学研究人员参考书，也可作为研究生培养的教学用书。

基于上述原因，特作序并向生命科学界同仁推荐此书。

中国工程院院士 夏家辉
中国医学遗传学国家重点实验室教授

2002年4月

前言

蛋白质的研究已有上百年的历史,但传统的蛋白质研究往往仅以某一个或一类蛋白质复合物为对象。直到1975年双向凝胶电泳技术的出现,从整体上对一个细胞甚至一种生物的全部蛋白质进行分析研究的想法才得以萌生。1982年Anderson提出了绘制人类蛋白质图谱的设想,1986年第一个蛋白质序列数据库SWISS-PROT在瑞士日内瓦大学建立。1990年人类基因组计划启动,而与此同时所取得的生物质谱技术的突破,推动了蛋白质大规模研究的开展,1994年澳大利亚学者Wilkins和Williams首先提出“蛋白质组”的概念。1997年出版了第一部蛋白质组学专著。2000年3月首次发表了一个生物体的完整蛋白质组。2001年2月和同年6月分别公布了人类基因组框架图和序列图谱,极大地促进了蛋白质组研究的兴起,成为后基因组计划的重要组成部分。

蛋白质组学作为后基因组时代研究的一个重要内容,将广泛深入生命科学和医药学的各个领域,其理论和技术的发展完善为困扰人类健康的一类重大疾病——肿瘤的研究带来了新的思维方式和研究领域,使人们从组织或细胞蛋白质整体水平这一全新角度来认识肿瘤成为可能。为此,我们广泛收集了当前国内外有关文献资料编写了《肿瘤蛋白质组学》一书,其框架为:绪论;“蛋白质样品制备与分离技术”(第一~第三章);“蛋白质组研究中的蛋白质鉴定与功能分析”(第四~第六章);“蛋白质组生物信息学”(第七~第九章);“肿瘤蛋白质组学研究”(第十~第十一章)。

尽管本书的编写者均是近年来从事蛋白质组学研究的专家教授和年轻学者,有一定的经验和体会,书稿经过反复讨论、修改、审校,主编也对全书作了必要的加工和润色,但限于编者水平,加之蛋白质组的概念诞生才短短几年,新方法、新技术接连产生,新观念、新思路不断涌现,因此欠缺或不妥之处在所难免,敬请广大读者批评指正,以便再版时不断充实和完善。

陈主初

2002年4月

目 录

绪 论

- 一、蛋白质组学的概念及其发展史 (1)
- 二、蛋白质组学技术方法概述 ... (3)
- 三、蛋白质组学在肿瘤研究中的应用 (4)

第一章 蛋白质样品制备技术

第一节 蛋白质制备常用技术 (8)

- 一、离液剂、还原剂及表面活性剂的选择 (8)
- 二、细胞破碎处理 (9)
- 三、组织细胞蛋白质样品的制备 ... (10)
- 四、细胞蛋白质样品的制备 (12)
- 五、体液蛋白质样品的制备 (15)

第二节 肿瘤样品蛋白质的抽提 ... (16)

- 一、肿瘤样品总蛋白质的抽提 ... (16)
- 二、肿瘤样品蛋白质的分级抽提 (19)
- 三、注意事项 (20)

第二章 蛋白质组双向凝胶电泳分离技术

第一节 双向聚丙烯酰胺凝胶电泳技术的发展 (24)

- 一、固相 pH 梯度等电聚焦双向电泳 (IPG-2DE) 技术 (24)
- 二、其他类型的 2-D 电泳技术 (25)
- 三、亚蛋白质组阵列重叠群技术 (25)

第二节 实验操作方法 (25)

- 一、蛋白质组双向凝胶电泳分析的原理 (25)
- 二、传统 IEF 方法 (ISO-DALT 电泳) (27)
- 三、固相 pH 梯度 - SDS 双向凝胶电泳 (IPG-DALT) 方法 (29)

第三章 双向凝胶电泳图像分析与数据库的构建

第一节 双向凝胶电泳数字化图像的获取 (37)

- 一、用 CCD 成像仪获取双向电泳凝胶图像 (38)
- 二、用扫描仪获取双向电泳凝胶图像 (38)

第二节 双向凝胶电泳图像分析 ... (38)

- 一、2-D 图像分析软件 PDQUEST 简介 (39)
- 二、2-D 图像分析软件 ImageMaster

简介	(43)
第三节 2-D 数据库的构建	(46)
一、用 Make2ddb 构建 2-D 数据库	(47)
二、用数据库软件构建 2-D 数据库	(48)
第四章 蛋白质鉴定技术	
第一节 概述	(49)
一、质谱技术	(49)
二、其他蛋白质鉴定技术	(51)
三、蛋白质组分析鉴定的高通量化	(52)
第二节 基质辅助激光解吸电离飞行时间质谱技术	(53)
一、基质辅助激光解吸电离技术 (MALDI)	(54)
二、飞行时间质谱 (TOF-MS)	(55)
三、MALDI-TOF 质谱测定肽质量指纹图 (PMF)	(57)
四、MALDI-TOF 质谱仪采用源后衰变 (PSD) 测定多肽序列	(62)
五、MALDI-TOF 质谱技术用于蛋白质 C-端序列分析	(64)
第三节 电喷雾质谱技术	(67)
一、电喷雾电离质谱的原理和特点	(67)
二、电喷雾质谱法测定蛋白质和多肽分子质量	(69)
三、电喷雾串联质谱法分析多肽和蛋白质的一级结构	(71)
四、电喷雾串联质谱法分析化学修饰多肽和蛋白质	(76)
五、液相色谱-电喷雾质谱法鉴定 2-D 胶蛋白质	(77)
第四节 微量蛋白质氨基酸组成分析	(84)
一、概述	(84)
二、Fmoc-Cl 柱前衍生及氨基酸组成分析	(85)
第五节 蛋白质与多肽化学法 N-端氨	

基酸序列测定	(92)
---------------------	------

一、Edman 降解的化学原理	(93)
二、Edman 降解过程中的起始收率和重复收率	(96)
三、PTH-氨基酸的高效液相色谱鉴定	(97)
四、蛋白质序列分析仪	(101)
五、Edman 降解序列测定中要注意的问题	(103)
六、双向电泳凝胶上蛋白质点的序列测定	(105)

第五章 定量蛋白质组学研究方法

第一节 同位素亲和标签方法	(110)
一、ICAT 方法的原理	(110)
二、蛋白质样品用 ICAT 试剂标记实验过程	(111)
三、ICAT 标记多肽用 LC-MS/MS 分析过程	(112)
四、注意事项	(112)
第二节 荧光差异显示双向凝胶电泳 (F-2D-DIGE) 方法	(113)
一、荧光差异双向电泳的实验原理	(113)
二、F-2D-DIGE 实验过程举例	(114)
三、注意事项及存在的问题	(115)

第六章 蛋白质与蛋白质相互作用的研究技术

第一节 酵母双杂交系统	(117)
一、概述	(117)
二、基本原理	(117)
三、方法	(119)
第二节 丝状噬菌体表面显示技术	(124)
一、丝状噬菌体的基本生物学特征	(124)
二、噬菌体表面显示的原理	(125)
三、噬菌体随机多肽文库	(126)

第三节 蛋白质芯片技术····· (131)	一、蛋白质分析工具软件····· (164)
一、蛋白质芯片种类····· (131)	二、蛋白质二级结构与三级结构分
二、蛋白质芯片探针的制备····· (132)	析····· (165)
三、蛋白质芯片的检测····· (133)	第十章 实体肿瘤的蛋白质组
四、蛋白质芯片技术的应用····· (133)	学研究
第七章 生物信息学与蛋白质	第一节 肿瘤蛋白质组学的研究策略
组信息学	和方法····· (168)
第一节 生物信息学概述····· (136)	一、肿瘤蛋白质组学的研究策略
一、生物信息学的定义····· (136)	····· (168)
二、生物信息学的诞生····· (136)	二、肿瘤蛋白质组学的研究方法
三、生物信息学的工作方式····· (137)	····· (170)
第二节 蛋白质组信息学概述····· (137)	第二节 实体肿瘤蛋白质组学研究的
一、蛋白质组信息学的兴起····· (137)	现状····· (171)
二、蛋白质组信息学的研究内容	一、鼻咽癌····· (171)
····· (138)	二、肺癌····· (171)
三、蛋白质组生物信息学展望·· (138)	三、乳腺癌····· (174)
第八章 蛋白质组研究常用的	四、肝癌····· (174)
生物信息资源	五、结肠癌····· (175)
第一节 蛋白质组研究信息资源概述	六、肾癌····· (175)
····· (140)	七、膀胱癌····· (176)
一、生物信息中心与数据库···· (140)	八、前列腺癌····· (176)
二、蛋白质芯片与蛋白质组学公司	九、其他肿瘤····· (177)
网址····· (145)	第十一章 血细胞和血液系统
第二节 蛋白质专家分析系统···· (145)	肿瘤的蛋白质组学
一、SWISS-PROT 的数据库格式	第一节 血细胞的蛋白质组学···· (180)
····· (146)	一、人淋巴细胞蛋白质组数据库
二、Prosite 数据库····· (150)	····· (180)
三、SWISS-2DPAGE 数据库·· (152)	二、人血小板胞浆蛋白质组学·· (181)
四、三维结构数据库····· (154)	第二节 部分血液系统疾病的蛋白质
五、ExPASy 中的链接····· (154)	组学····· (187)
第九章 生物信息学软件在蛋	一、Burkitt 淋巴瘤的蛋白质组研究
白质组研究中的运用	····· (187)
第一节 常用的蛋白质组生物信息学	二、2-DE 技术揭示急性白血病与
软件····· (158)	癌性蛋白 18 磷酸化高度相关
一、双向凝胶电泳分析软件···· (158)	····· (189)
二、质谱结果的数据库查询···· (159)	三、维甲酸诱导急性早幼粒细胞白
三、其他工具软件····· (160)	血病细胞凋亡始动时相的蛋白
第二节 蛋白质结构分析软件使用简	质组学分析····· (189)
介····· (163)	

绪 论

人类基因组序列图谱初稿的公布,在揭示基因组的精细结构的同时,也凸现出基因数量有限性和基因结构的相对稳定性,这与生命现象的复杂性和多变性之间存在着巨大反差。这种反差促使人们认识到:基因只是遗传信息的载体。要研究生命现象,阐释生命活动的规律,只了解基因组的结构是远远不够的,这也使得人们对于生命活动的直接执行者——蛋白质的重要性有了更深刻的理解,因此,对蛋白质的数量、结构、性质、相互关系和生物学功能进行全面和深入的研究,成为生命科学研究的迫切需要和重要任务。一个以“蛋白质组”(proteome)为研究重点的生命科学新时代已悄然到来。

一、蛋白质组学的概念及其发展史

随着后基因组时代的到来,蛋白质组研究越来越受到国内外科学工作者的密切关注,并以其特有的思维方法和技术手段在解决生物学重大问题上开始显示出强大威力。可以相信,随着蛋白质组研究的不断深入,它在揭示生长、发育、凋亡、分化、信号传导和代谢调控等生命活动的规律上将会有新突破,对探讨重大疾病的发生机制、疾病的诊断防治和新药开发将提供重要的理论基础。

(一) 蛋白质组学的概念

“蛋白质组”一词的英文是 PROTEOME,它由 PROTEins 和 genOME 两个词组合而成,意思是 proteins expressed by a genome,即基因组表达的蛋白质。广义上讲,蛋白质组(proteome)是指“一个细胞或一个组织基因组所表达的全部蛋白质”。它是对应于一个基因组的所有蛋白质构成的整体,而不是局限于一个或几个蛋白质。由于同一基因组在不同细胞、不同组织中的表达情况各不相同,即使是同一细胞,在不同的发育阶段、不同的生理条件甚至不同的环境影响下,其蛋白质的存在状态也不尽相同。因此,蛋白质组是一个在空间和时间上动态变化着的整体。

蛋白质组学(proteomics)是指应用各种技术手段来研究蛋白质组的一门新兴学科,其目的是从整体的角度分析细胞内动态变化的蛋白质组成成分、表达水平与修饰状态,了解蛋白质之间的相互作用与联系,揭示蛋白质功能与细胞生命活动规律。要对“全部蛋白质”进行研究是非常困难的,“功能蛋白质组学(functional proteomics)”的提出解决了这一难题,其研究对象是功能蛋白质组,即细胞在一定阶段或与某一生理现象相关的所有蛋白质,它是介于对个别蛋白质的传统蛋白质化学研究和以全部蛋白质为研究对象的蛋白质组学研究之间的层次,并把目标定位在蛋白质群体上,这一群体可大可小,从局部入手研究蛋白质组的各个功能亚群体,以便将来把多个亚群体组合起来,逐步描绘出接近于生命细胞的“全部蛋白质”的蛋白质组图谱。功能蛋白质组学的研究为实际运作带来方便,人们可利用现有的技术手段来研究蛋白质组这一极限群体中的各蛋白质功能亚群,从理论和技术上使以“全部蛋白质”为研究对象的蛋白质组学这一抽象概念具体化,使研究者们更易于从时、空、量效方面动态、整

体、深入地研究生理状态下同一组织细胞在不同发育阶段或同一组织细胞在不同个体间或同一基因组在不同组织细胞间,以及病理情况下同一疾病的不同发展阶段的蛋白质表达模式和功能模式的变化,揭示一些重要的生命现象和一些重大疾病的发生发展规律。狭义上讲,蛋白质组是指不同条件下细胞内蛋白质的变化,比如正常细胞和异常细胞之间,细胞用药和不用药之间的蛋白质表达谱的差别,这在疾病研究和药物筛选上很有意义,是目前蛋白质组学在应用研究方面最具前景的领域。当然,随着蛋白质组研究的不断发展,蛋白质组学的概念也将不断发展和深化。

(二) 蛋白质组学的产生与发展

20 世纪中期以来,随着 DNA 双螺旋结构的提出和蛋白质空间结构的 X 射线解析,开始了分子生物学时代,对遗传信息载体 DNA 和生命功能的主要体现者蛋白质的研究,成为生命科学研究的主要内容。20 世纪 90 年代初期,美国生物学家提出并实施了人类基因组计划,经过各国科学家多年的努力,人类基因组计划取得了巨大成绩,一些低等生物的 DNA 序列已被阐明,人类 DNA 序列的框架图已经测定,迄今已测定的表达序列标签(EST)几乎覆盖了人类所有基因。在这样的形势下,生命科学已进入了后基因组时代。在后基因组时代,生物学研究的重点已从揭示生命所有遗传信息转移到在整体水平上对生物功能的研究。这种转向的第一个标志就是产生了一门称为功能基因组学的新学科。功能基因组学的主要任务是解析和综合大量的遗传信息,阐明基因遗传信息与人类生命活动之间的联系。基因的功能是通过其产物 mRNA 和蛋白质来体现的。近年来利用基因表达连续分析(SAGE)、微阵列(microarray)和 DNA 芯片(chip)等新技术研究 mRNA 水平上的基因活动规律取得了较大进展。但 mRNA 水平的基因表达状况并不能完全代表蛋白质水平的状况, mRNA 与蛋白质间的相关系数仅为 0.4~0.5,蛋白质才是生命功能的主要执行者,由于它存在着翻译后的加工修饰、转移定位、构象变化、蛋白质与蛋白质及蛋白质与其他生物大分子相互作用等自身特点,因此难以从 DNA 和 mRNA 水平得到解答,从而促使人们从组织或细胞内整体蛋白质的组成、表达和功能模式去研究生命活动的基本规律。自 1994 年澳大利亚学者 Williams 和 Wilkins 首先提出与基因组相对应的“蛋白质组”概念,并开始从整体蛋白质水平研究生命现象以来,蛋白质组研究在国际上进展十分迅速,不论是基础理论,还是技术方法,都在不断地进步和完善。许多种细胞的蛋白质组数据库已经建立,相应的国际互联网站也层出不穷,众多的制药厂和公司在巨大财力的支持和商业利益的诱惑下纷纷加入蛋白质组研究领域,蛋白质组研究的文章每年成倍增长。1996 年澳大利亚建立了世界上第一个蛋白质组研究中心(Australia proteome analysis facility, APAF)。在政府部门的大力支持下,丹麦、加拿大也先后成立了蛋白质组研究中心。1997 年召开了第一次国际“蛋白质组学”会议,预测 21 世纪生命科学的重心将从基因组学转移到蛋白质组学,为生命科学和医药学领域的研究带来了新的生机。1998 年在美国旧金山召开了第二届国际蛋白质组学会议,参加者大都来自各大药厂和公司。1999 年 1 月在英国伦敦举行了应用蛋白质组会议。1999 年 5 月在日本召开的国际电泳会议上,约 1/3 的文章与蛋白质组有关。目前,中国国家自然科学基金委员会关于蛋白质组重大项目的研究已启动,肿瘤蛋白质组研究也已列入我国 973 和 863 项目。中国科学院上海生物化学研究所、中国军事医学科学院、湖南师范大学、中南大学湘雅医学院等单位相继开展了蛋白质组研究。1998 年在中国科学院上海生物化学研究所举办了两次全国性的蛋白质组学研讨会,并在此基础上,提出了功能蛋白质组学的研究战略。可以相信在不久的将来,“人类蛋白质组计划”必将启动,而且规模比“人类基因组计划”更大,产生的影响更久远,将是继基因

组研究之后的又一“大科学”。

二、蛋白质组学技术方法概述

蛋白质组学主要涉及有两方面的内容：一是研究蛋白质组的组成成分，即蛋白质组表达模式的研究；二是研究蛋白质组的功能，即蛋白质组功能模式的研究，目前主要集中在蛋白质组表达模式方面。蛋白质组表达模式研究的支撑技术主要有双向凝胶电泳和以质谱为代表的蛋白质鉴定技术及生物信息学。

双向凝胶电泳（2-DE）技术于1975年由O'Farrell等创立，其原理是第一向在高压电场下对蛋白质进行等电聚焦（IEF），再在第一向垂直方向上进行第二向SDS-聚丙烯酰胺凝胶电泳（SDS-PAGE）。等电聚焦电泳由最初的载体两性电解质管胶电泳发展到目前的固相pH梯度（IPG）凝胶电泳，由于采用了IPG胶条从而避免了因载体两性电解质引起的聚焦时间延长、pH梯度不稳定、阴极漂移等现象。目前IPG双向凝胶电泳的分辨率达到了1万多个蛋白质点，但对过于偏酸或偏碱、高分子质量、极微量蛋白质以及难溶性蛋白质的分辨仍感困难。由于双向凝胶电泳对批量蛋白质可实现一次性分离，具有高灵敏度和高分辨率、便于计算机进行图像分析处理、可以很好地与质谱分析等鉴定方法匹配的优点，因而成为目前分离蛋白质组分的核心技术。此外，双向高效柱层析、毛细管电泳也是分离蛋白质组分的有效技术。

对分离的蛋白质进行鉴定是蛋白质表达模式研究的又一重要内容，通常采用的有蛋白质微量测序、氨基酸组成分析等传统的蛋白质鉴定方法，但这些方法费时、费力、不易实现高通量分析。因此一种新的蛋白质鉴定技术——质谱（MS）法受到了人们的重视和应用，其基本原理是样品分子离子化后，根据离子间质荷比（ m/z ）的差异来分离并确定样品的分子质量。质谱在20世纪初就已经产生，多用于无机物或小分子有机物的鉴定，直到80年代末随着“软电离”技术的出现而进入生物大分子（如蛋白质）的鉴定领域。所谓“软电离”是指样品分子电离时保留整个分子的完整性，不会形成碎片离子。软电离质谱用于大分子的鉴定，主要有以下特点：灵敏度高、快速、能同时提供样品的精确分子质量和结构信息、既可定性又可定量、并能有效地与各种色谱联用，如用HPLC/MS来分析复杂体系。目前用于蛋白质鉴定的质谱主要有两种：电喷雾质谱（ESI-MS）和基质辅助激光解吸/电离飞行时间质谱（MALDI-TOF-MS）。但要精确鉴定某一蛋白质通常还得联合几种鉴定技术。最近，国外建立了一种将MALDI-MS技术直接应用于组织切片或印片的方法（Pierre C et al, 1999），即组织切片或印片原位质谱分析技术，取得了较满意的结果。

生物信息学是随着人类基因组计划、计算机技术、网络技术的发展而诞生的一门新兴学科，是蛋白质组学的一个重要的技术平台。生物信息学研究生物信息的采集、加工、分析、存储、传播等各个方面，它通过综合应用数学、计算机科学、工程科学以及生物学的技术、方法分析大量而复杂的生物学数据来揭示生物学的奥秘。生物信息学是蛋白质组学研究的一个不可缺少的部分，其在蛋白质组学中的研究有两个重要应用：一是构建和分析双向凝胶电泳图谱，二是数据库的搜索与构建。目前应用最普遍的数据库是NRDB和dbEST数据库。NRDB由SWISS-PROT和GENPETP等几个数据库组成。dbEST是由美国国家生物技术信息中心/欧洲生物信息学研究所（National Centre for Biotechnology Information/European Bioinformatics Institute, NCBI/EBI）共同编辑的核酸数据库，包括许多生物体的表达序列标签（EST）。用肽序列标签（PST）或部分序列信息最适合查寻dbEST数据库。最近有报道，强调用蛋白质质

谱分析数据查寻 EST 数据库以鉴定研究者最感兴趣的未知蛋白质的重要性,表明人类 EST 数据库可满足用质谱数据快速识别哺乳动物多蛋白质复合物的要求。

蛋白质功能模式的研究是蛋白质组研究的最终目标,其主要研究目标是要揭示蛋白质组成员间的相互作用、相互协调的关系,并深入了解蛋白质的结构与功能的相互关系,以及基因结构与蛋白质结构功能的关系。目前,蛋白质功能模式研究的主要技术有酵母双杂交系统、噬菌体展示、生物传感芯片质谱、基因工程和蛋白质工程中的突变表达分析、磁共振成像、X 线晶体衍射分析、蛋白质芯片技术等。

三、蛋白质组学在肿瘤研究中的应用

运用蛋白质组学研究手段,通过比较正常和病理情况下细胞或组织中蛋白质在表达数量、表达位置和修饰状态上的差异,可以发现与病理改变有关的蛋白质和疾病特异性蛋白质,这些蛋白质既可为疾病发病机制提供线索,也可作为疾病诊断的分子标记,还可作为治疗和药物开发的靶标,其中肿瘤是蛋白质组研究应用最为广泛的领域。肿瘤是困扰人类的一类重大疾病,它的发生是机体正常细胞在多因素、多基因作用下经过多途径而发生转化,失去原有的生长分化调节而异常增生的结果。在这一癌变过程中,不论是环境因素还是遗传因素最终都是通过基因及其相应的蛋白质来发挥作用。在人类基因组计划(HGP)和癌基因组解剖计划(CGAP)推动下,人们从基因水平对肿瘤进行了广泛的研究,先后发现了癌基因、抑癌基因和 DNA 错配修复基因等诸多肿瘤相关基因,但基因的功能活动最终要靠蛋白质来体现。已知在 DNA→mRNA→蛋白质的过程中存在着转录后剪切和翻译后的修饰加工及蛋白质合成后的转移定位等多种变化,所以 DNA 或 mRNA 水平、状况并不完全代表蛋白质的水平、状况,也就是说基因的种类和数目不等于 mRNA 的种类和数目, mRNA 的种类和数目也不等于蛋白质的种类和数目。因此有必要从蛋白质整体水平来揭示肿瘤的发生发展,阐明其癌变本质。肿瘤细胞是由正常细胞转变而来,与其起源的正常细胞相比,在蛋白质上既有所同也有所异。它不仅拥有一些正常细胞所没有的某些特异蛋白质或缺少正常细胞所具有的某些特异蛋白质,而且也有许多与正常细胞相同的蛋白质,只不过这些蛋白质可能在表达水平或翻译后的修饰加工上与正常细胞存在差别。蛋白质组学概念的提出及理论、技术的逐步发展和完善,使人们从组织或细胞蛋白质整体水平这一全新角度来研究人类肿瘤变为现实,并成为肿瘤研究的一个新热点。比较患同一肿瘤的不同个体的肿瘤组织(细胞)之间或同一个体肿瘤细胞与正常起源细胞之间蛋白质组的异同,可以发现一些肿瘤特异性的蛋白质,深入了解这些肿瘤特异性蛋白质的结构和功能,将极大地促进肿瘤的发病机制、诊断、治疗和新药开发的研究,具有重要的理论和实际意义。

肿瘤蛋白质组学研究近年已开始启动,如对鼻咽癌、膀胱癌、乳腺癌、肾癌、肺癌、卵巢癌、前列腺癌、肝癌、结直肠癌、神经母细胞瘤等多种实体肿瘤和血液系统肿瘤,国内外均进行了一定的研究,并取得了一些有意义的结果。

在实体肿瘤方面,中南大学湘雅医学院肿瘤研究所与湖南师范大学生命科学院率先在国内外采用双向凝胶电泳和质谱技术分析了人鼻咽癌细胞系 HONE1 的总蛋白质和人肺鳞癌细胞系 NCI-H520 的蛋白质组成分及人肺腺癌细胞系 A-549 与永生化人正常支气管上皮细胞系 HBE 间差异表达的蛋白质。中国科学院上海生化所采用双向凝胶电泳和液相色谱-离子阱质谱联用技术分析了人肝癌细胞系 BEL-7404 和人正常肝细胞系 L-02 间差异表达的蛋白质,以及反义表皮生长因子受体序列(antisense epidermal growth factor receptor sequence)转染

前后的人肝癌细胞系 (JX-0 和 JX-1) 的蛋白质组改变情况。日本 Okuzama 和 Hirano 等利用非酶法样品制备技术、2-DE 和质谱 (ESI-MS 和 MALDI-MS) 分析、结合 SWISS-PROT 数据库查询, 也分离了十几种与肺癌组织学类型和分化程度相关的蛋白质和多肽, 其中 TAO1 (35kDa、pI5.45) 和 TAO2 (35kDa、pI5.29) 两种多肽的表达水平不仅反应原发性肺腺癌的分化程度, 而且是区分原发性和转移性肺腺癌的价值指标。丹麦 Celis 等采用双向电泳、微量蛋白质测序、质谱分析、冰冻切片免疫荧光检测等方法对膀胱癌组织和原代培养纯化的膀胱癌细胞进行了膀胱癌蛋白质组分析, 发现了二十多种膀胱癌相关蛋白质, 建立了膀胱癌蛋白质 2-DE 数据库。瑞典 Franzen 等、澳大利亚 Rasmussen 等和 Williams 等应用 2-DE、免疫印迹分析法、Edman 降解测序、反向 HPLC、毛细管电泳、电喷雾质谱和生物信息等方法分别鉴定了乳腺癌和乳腺癌细胞系的几十种相关蛋白质, 在这些相关蛋白质中, 大多数蛋白质表达改变是由于翻译后修饰造成的, RNA 水平的变化未能探测到, 其中有些蛋白质是以前尚未报道的。意大利 Sarto 等通过比较分析正常肾组织和肾细胞癌的蛋白质组双向电泳图谱, 发现四种蛋白质在肾细胞癌中缺失, 其中两个与线粒体氧化还原功能有关, 说明线粒体氧化还原功能障碍在肾细胞癌的发生、发展中起重要作用; 通过生物信息分析进一步发现肾细胞癌蛋白质表达变化与其染色体变化存在着相关性, 并可作为肾癌的诊断标志物。瑞典 Alaiya 等从蛋白质组出发分析了卵巢癌、良性卵巢瘤及癌旁组织的蛋白质表达谱, 发现 9 个多肽标志物表达的变化较大, 联合应用这 9 个多肽标志物进行相似性分析发现, 在卵巢癌有 5 个以上的多肽标志物表达, 在良性卵巢瘤则不多于 3 个多肽, 癌旁组织有 0~6 个不等, 因此对 2-DE 分离的多肽进行相似性分析对卵巢癌的基因表达特征和诊断研究均是有用的。

蛋白质组研究方法也成功地被用来探测和分析化学致癌物诱发肿瘤的相关蛋白质及其特征。德国 Zeindl-Eberhart 等通过氨基酸组成分析和微量蛋白质末端测序, 在化学致癌物诱发的鼠肝癌中发现分子量相同 (33~35KDa)、等电点 (pI) 不同的 8 种蛋白质的变化形式为醛糖还原酶的同分异构体, 其氨基酸序列与已知的晶体醛糖还原酶有 98.5% 的同源。肝癌源性的醛糖还原酶只在肝癌组织和胚胎肝表达, 目前在人肝细胞癌中发现一种与鼠肝癌源性的醛糖还原酶类似的蛋白质。

自 1987 年 Tsuchiya 等应用 2-DE 分析急性淋巴细胞白血病细胞的蛋白质表达谱以来, 很多学者加入了血液系统肿瘤的蛋白质组学研究行列, 研究资料不断累积。Hanash 等通过多年的研究, 建立了淋巴组织 (包括正常淋巴细胞、淋巴系统增生性疾病和白血病) 蛋白质的 2-DE 数据库, 该数据库庞大, 含 9167 个 2-DE 图谱。Müller 等应用 2-DE 和电喷雾串联质谱技术, 不仅建立了人 Burkitt 淋巴瘤 BL-60 细胞蛋白质的 2-DE 数据库, 而且成功地鉴定了 33 种 BL-60 相关蛋白质, 其中 3 种为未知蛋白质。Voss 等采用 2-DE 技术研究了慢性 B 淋巴细胞白血病病人临床特征与 B 淋巴母细胞蛋白质表达谱的关系, 结果显示, 病人临床特征与 B 淋巴母细胞的蛋白质表达谱密切相关, 生存期短的病人, 癌细胞的氧化还原酶、热休克蛋白 27 和二硫化物同分异构酶的表达水平发生明显变化, 并认为这些蛋白质可能与耐药有关。另外, 我国军事医学科学院放射医学研究所对全反式维甲酸 (ATRA) 诱导人急性早幼粒白血病细胞株 HL-60 细胞凋亡的蛋白质表达谱进行了分析, 并着重研究了凋亡启动时相的蛋白质组学变化。

如上所述, 肿瘤蛋白质组学研究已初见端倪, 其研究成果将为建立肿瘤癌变各阶段的蛋白质组数据库、发现肿瘤相关蛋白质、分析肿瘤相关蛋白质结构和功能以及揭示肿瘤发生的分子机制提供依据, 为寻找肿瘤标志物、特异的治疗肿瘤的靶分子和开发新的抗肿瘤药物提

供重要的、直接的线索,并通过合成相应的蛋白质和抗体或制备蛋白质芯片为肿瘤诊治服务等。可以相信,蛋白质组学将在人类攻克癌症的事业中发挥出越来越重要的作用。

(陈主初)

参考文献

- 1 李峰,关勇军,陈主初,等.人鼻咽癌细胞系 HONE1 的蛋白质双向凝胶电泳分析.中国癌症杂志, 2001, 11(4): 294 ~ 298
- 2 詹显全,陈主初,关勇军,等.双向凝胶电泳-飞行时间质谱法分析人肺鳞癌细胞 NCI-H520 蛋白质组分.癌症, 2001, 20(6): 575 ~ 582
- 3 詹显全,关勇军,陈主初,等.人肺腺癌细胞 A-549 和正常细胞 HBE 的蛋白质组差异分析.生物化学与生物物理学报, 2002, 34(1): 50 ~ 56
- 4 Alaiya A, Franzen B, Auer G, et al. Cancer proteomics: from identification of novel markers to creation of artificial learning models for tumor classification. *Electrophoresis*, 2000, 21: 1210 ~ 1212
- 5 Alaiya A, Franzen B, Fujioka K, et al. Phenotypic analysis of ovarian carcinoma: polypeptide expression in benign, borderline and malignant tumors. *Int J Cancer*, 1997, 73(5): 678 ~ 683
- 6 Anderson N. Proteome and proteomics: new technologies, new concepts, and new words. *Electrophoresis*, 1998, 19: 1853 ~ 1861
- 7 Celis J, Celis P, Ostergaard M, et al. Proteomics and immunohistochemistry define some of the steps involved in the squamous differentiation of the bladder transitional epithelium: a novel strategy for identifying metaplastic lesions. *Cancer Res*, 1999, 59(12): 3003 ~ 3009
- 8 Celis J, Wolf H, Ostergaard M. Bladder squamous cell carcinoma biomarkers derived from proteomics. *Electrophoresis*, 2000, 21(11): 2115 ~ 2121
- 9 Cordwell S, Wilkins M, Cerpa-Poljak A, et al. Cross-species identification of proteins separated by two-dimensional gel electrophoresis using matrix-assisted laser desorption/ionization/time-of-flight mass spectrometry and amino acid composition. *Electrophoresis*, 1995, 16(3): 438 ~ 443
- 10 Franzen B, Linder S, Alaiya A, et al. Analysis of polypeptide expression in benign and malignant human breast lesions: down-regulation of cytokeratins. *Br J Cancer*, 1996, 74(10): 1632 ~ 1638
- 11 Franzen B, Linder S, Alaiya A, et al. Analysis of polypeptide expression in benign and malignant human breast lesions. *Electrophoresis*, 1997, 18(3 ~ 4): 582 ~ 587
- 12 Hanash S, Teichrow D. Mining the human proteome: experience with the human lymphoid protein database. *Electrophoresis*, 1998, 19(11): 2004 ~ 2009
- 13 Hirano T, Fujioka K, Franzen B, et al. Relationship between TAO1 and TAO2 polypeptides associated with lung adenocarcinoma and histocytological features. *Br J Cancer*, 1997, 75(7): 978 ~ 985
- 14 Jain K. Applications of proteomics in oncology. *Pharmacogenomics*, 2000, 1(4): 385 ~ 393
- 15 Jungblut PR, Zimny AU, Zeindl EE, et al. Proteomics in human disease: cancer, heart and infectious diseases. *Electrophoresis*, 1999, 20(10): 2100 ~ 2110
- 16 Müller E, Schumann M, Rickers A, et al. Study of Burkitt lymphoma cell line proteins by high resolution two-dimensional gel electrophoresis and nanoelectrospray mass spectrometry. *Electrophoresis*, 1999, 20: 320 ~ 333
- 17 Okuzawa K, Franzen B, Lindholm J, et al. Characterization of gene expression in clinical lung cancer materials by two-dimensional polyacrylamide gel electrophoresis. *Electrophoresis*, 1994, 15(3 ~ 4): 382 ~ 390
- 18 Pandey A, Mann M. Proteomics to study genes and genomes. *Nature*, 2000, 405: 837 ~ 846
- 19 Pierre C, Markus S, Richard MC. Direct profiling of proteins in biological tissue sections by MALDI mass spectrometry.

- try. *Anal Chem*, 1999, 71: 5 263 ~ 5 270
- 20 Rasmussen RK, Ji H, Eddes JS, et al. Two-dimensional electrophoretic analysis of human breast carcinoma proteins: mapping of proteins that bind to the SH3 domain of mixed lineage kinase MLK2. *Electrophoresis*, 1997, 18(3 ~ 4): 588 ~ 598
 - 21 Richard M, Terry B, and Jocelyn G. Molecular imaging of biological samples: localization of peptides and proteins using MAILDI-TOF MS. *Anal Chem*, 1997, 69: 4 751 ~ 4 760
 - 22 Sarto C, Frutiger S, Cappellano F, et al. Modified expression of plasma glutathione peroxidase and manganese superoxide dismutase in human renal cell carcinoma. *Electrophoresis*, 1999, 20(17): 3 458 ~ 3 466
 - 23 Sarto C, Marocchi A, Sanchez J, et al. Renal cell carcinoma and normal kidney protein expression. *Electrophoresis*, 1997, 18(3 ~ 4): 599 ~ 604
 - 24 Tsuchiya H, Migita M, Nunoi H, et al. Analysis of cell proteins in lymphoblasts of acute lymphocytic leukemia by two-dimensional gel electrophoresis. *Acta Haematol*, 1987, 77(2): 65 ~ 71
 - 25 Voss T, Ahorn H, Haberl P, et al. Correlation of clinical data with proteomics profiles in 24 patients with B-cell chronic lymphocytic leukemia. *Int J Cancer*, 2001, 91 (2): 180 ~ 186
 - 26 Walsh BJ, Molloy MP, Williams KL. The Australian Proteome Analysis Facility (APAF): Assembling large scale proteomics through integration and automation. *Electrophoresis*, 1998, 19: 1 883 ~ 1 890
 - 27 Wan JH, Wang JL, Cheng HP, et al. Proteomic analysis of apoptosis initiation Induced by all-trans retinoic acid in human acute promyelocytic leukemia cells. *Electrophoresis*. 2001, 22: 3 026 ~ 3 037
 - 28 Williams K, Chubb C, Huberman E, et al. Analysis of differential protein expression in normal and neoplastic human breast epithelial cell lines. *Electrophoresis* 1998, 19 (2): 333 ~ 343
 - 29 Wilkins MR, Williams KL, Hohstrasser DF, et al. (ed) *Proteome Research: New Frontiers in Fuctional Genomics*. New York: Springer, 1997
 - 30 Yu L, Shao X, Jiang W, et al. Proteome alterations in human hepatoma cells transfected with antisense epidermal growth factor receptor sequence. *Electrophoresis*, 2001, 22: 3 001 ~ 3 008
 - 31 Yu L, Zeng R, Shao X, et al. Identification of differentially expressed proteins between human hepatoma and normal liver cell lines by two-dimensional electrophoresis and liquid chromatography-ion trap mass spectrometry. *Electrophoresis*, 2000, 21: 3 058 ~ 3 068
 - 32 Zeindl-Eberhart E, Jungblut P, Rabes HM. Expression of tumor-associated protein variants in chemically induced rat hepatomas and transformed rat liver cell lines determined by two-dimensional electrophoresis. *Electrophoresis*, 1994, 15 (3 ~ 4): 372 ~ 381
 - 33 Zeindl-Eberhart E, Jungblut PR, Otto A, et al. Further characterization of a rat hepatoma-derived aldose-reductase-like protein-organ distribution and modulation in vitro. *Eur J Biochem*, 1997, 247 (3): 792 ~ 800

第一章 蛋白质样品制备技术

蛋白质组是指某种细胞或组织中基因组表达的全部蛋白质。实践中更多的情况是研究功能蛋白质组，即细胞在某一特定阶段或某一生理状态下基因表达的所有蛋白质。由于蛋白质组研究是对不同时间和不同空间发挥功能的蛋白质整体的研究，因此如何从细胞组织中尽可能完整地将蛋白质以溶解状态提取出来，是蛋白质组研究的首要步骤。

第一节 蛋白质制备常用技术

在蛋白质制备中主要包括细胞组织的破碎裂解；蛋白质增溶溶解以破坏蛋白质与蛋白质分子之间，蛋白质与非蛋白质之间的共价与非共价相互作用；变性及还原；去除非蛋白质组分，如核酸、脂类等。为了达到这一目的，在蛋白质制备过程中需使用表面活性剂、还原剂及离液剂。通常使用的蛋白质样品溶解液或细胞裂解缓冲液含有 8 ~ 9.8 mol/L 脲、蛋白酶抑制剂、50 ~ 100 mmol/L DTT、4% CHAPS、40 mmol/L Tris 和 0.5% Pharmalyte pH 3 ~ 10。但样品中各种蛋白质的溶解度差别很大，使用这种标准提取液并不能真正做到将所有蛋白质提取出来。某些蛋白质如膜蛋白、碱性蛋白如果使用一般的溶解液则溶解性较差，样品溶解不好就会减少分离得到的蛋白质数量，甚至造成在等电聚焦时某些蛋白质产生沉淀，并影响从第一向转移到第二向的蛋白质数量。因此，为了提高总蛋白质的溶解性，常常要使用一些新的表面活性剂、还原剂及离液剂以增加蛋白质的溶解性。由于不同的蛋白质溶解性不同及在细胞中存在部位的差别，所以有时需要采取分步提取的方法或进行亚蛋白质组分的顺序分步提取，才能将细胞中的全部蛋白质分离提取出来。样品的制备没有一种通用方法，不同来源的样品需要不同的提取和裂解技术。但在蛋白质抽提过程中，有几条共同的原则需要遵循：一是尽可能溶解全部蛋白质，打断蛋白质之间的非共价键结合，使样品中的蛋白质以分离的多肽链形式存在；二是避免蛋白质的修饰作用和蛋白质的降解作用；三是避免脂类、核酸、盐等物质的干扰作用；四是蛋白质样品与第一向电泳的相容性。

一、离液剂、还原剂及表面活性剂的选择

(一) 离液剂 (chaotropes)

最普遍的离液剂是脲，脲的作用主要是改变或破坏氢键等次级键的结构，使蛋白质的肽链伸展开来，充分暴露疏水中心，降低接近疏水残基的能量域，但通常当脲和表面活性剂 CHAPS 联合使用时，会使某些蛋白质吸附在固相 pH 梯度凝胶中而丢失。因此，近年常将一种新的离液剂硫脲和脲联合使用，以增加蛋白质在 IPG 胶中的溶解性。硫脲的使用大大改善了蛋白质的溶解性能，特别是改善了膜蛋白的提取。但硫脲在水中溶解性差，需用高浓度的脲助溶，因此在实际应用中常用 2 mol/L 硫脲与 5 ~ 7 mol/L 脲联合使用，以利于膜蛋白的提取。

(二) 表面活性剂

蛋白质在去折叠后, 会暴露出大量疏水性残基, 因此常需使用表面活性剂破坏蛋白质分子之间的疏水相互作用。常使用的表面活性剂有阴离子去污剂 SDS, 非离子去污剂 Triton X-100 和 NP-40, 两性离子去污剂 CHAPS 等。离子型去污剂如 SDS 不利于稳定等电聚焦中蛋白质的等电点, 因此一般不宜在等电聚焦样品溶解液中使用。但 SDS 能有力地溶解膜蛋白, 故可用于样品处理的初始阶段, 当其浓度低于 0.25% 时, 对等电聚焦不会产生太大的影响。目前可用于 IEF 的表面活性剂主要为 CHAPS 等非离子或两性离子表面活性剂。在高浓度的增溶剂里特别是含高浓度的硫脲时, 这些表面活性剂溶解蛋白质的能力仍然有限。含有长线性烷基尾端的 SB3-10 等两性离子表面活性剂比 CHAPS 更有效, 但 SB3-10 不溶于浓度高于 5mol/L 的脲中。Rabilloud (1996, 1998) 提出以 5mol/L 脲、2mol/L 硫脲、2% CHAPS 及 2% SB3-10 来溶解需要强烈表面活性剂的蛋白质, 用 7mol/L 脲、2mol/L 硫脲、4% CHAPS 用于需要高浓度离液剂的蛋白质。为了解决表面活性剂与离液剂在高浓度脲中的溶解度低问题, 有人合成了一种新的包括具有线性碱性尾的酰胺硫代甜菜碱 (ASB) 或具有 alkyl-aryl 尾的 ASB 如 C8 以增加其在高浓度脲中的溶解性。由于它们比 SB3-10 具有更多的极性头, 能溶解于高浓度的脲中。因 ASB 及 C8 比 CHAPS 具有更强的膜蛋白溶解力, 所以更适合于对膜蛋白的提取。

(三) 还原剂

在蛋白质制备中常需要断裂蛋白质分子中的二硫键, 以利于肽链的分离, 因此样品还原的好坏也会影响到蛋白质的分离效果。一般使用 β -巯基乙醇和二硫苏糖醇 (DTT) 作还原剂, 但 DTT 本身带有电荷, 在等电聚焦时, 常常会迁移到 pH 范围以外, 从而使某些蛋白质的二硫键重新配对, 使溶解度降低而重新沉淀下来。采用非离子型还原剂如三丁基膦 (tributyl-phosphine, TBP) 可大大增加蛋白质的溶解性 (Herbert B.R. et al 1998), 并可帮助蛋白质从第一向转移到第二向。

二、细胞破碎处理

细胞破碎主要采用机械和化学方法进行 (Gull M, 1990)。由于在细胞破碎时会释放出蛋白酶, 所以应加蛋白酶抑制剂。为了使蛋白酶及其他引起蛋白质修饰的酶迅速失活, 一般尽量将样品直接加入酶的变性裂解液中, 并在低温及使用预先冷冻的溶液下进行操作。常用的细胞破碎方法分为以下两种:

(一) 轻微裂解方法

(1) 渗透溶胞法: 这是一种非常温和的裂解细胞的方法, 可用于进一步进行细胞分级为亚细胞组分, 如血细胞及组织培养细胞悬浮于渗透溶胞溶液中, 细胞溶胀而释放出细胞内容物。

(2) 冻融裂解: 可将细菌细胞及组织培养细胞放在液氮中冻结后, 迅速融解, 再冻结, 再融解, 如此循环多次。

(3) 去污剂裂解: 通过含去污剂的裂解液处理组织、细胞以裂解细胞膜, 使细胞内含物释放出来。一般常将细胞直接置裂解液或样品液中, 进行悬浮处理。

(二) 更强烈的细胞破碎方法

当破碎不容易破碎的细胞或固体组织中的细胞时, 因有坚韧细胞膜, 需要采用激烈的方法进行细胞破碎, 这些方法可使细胞完全破碎裂解, 在操作过程中应避免产热及产生泡沫, 以免引起蛋白质变性。主要的方法有: