

现代生物技术前沿

〔美〕D. C. 利布莱尔 著

张继仁 译

高友鹤 校

蛋白质组学导论

——生物学的新工具



科学出版社

www.sciencep.com

Q51

LBL

C1 现代生物技术前沿

〔美〕D. C. 利布莱尔 著

张继仁 译

高友鹤 校

蛋白质组学导论

——生物学的新工具



科学出版社
北京

图字: 01-2004-2025

内 容 简 介

本书介绍了分析蛋白质和肽的各种方法,重点阐述了不同质谱仪及相关数据库检索算法的基本原理和使用方法,详细描述了质谱在蛋白质组学中的应用。蛋白质组学领域最权威的科学家之一 John R. Yates 教授称本书是蛋白质组学的极好的导论和综述,可供有生物化学背景的学生和科学家使用,书写流畅,深入浅出。

本书可用作从事生命科学和医学研究的专业人员的参考书,也可用作学习生命科学和生物技术的本科生和研究生的教材。

The original English language work has been published by HUMANA PRESS,
Totowa, New Jersey, U. S. A.

©2002 by Humana Press. All rights reserved.

图书在版编目 (CIP) 数据

蛋白质组学导论:生物学的新工具/(美)利布莱尔(D. C. Liebler)
著. 张继仁译. —北京:科学出版社, 2005

(现代生物技术前沿)

ISBN 7-03-014258-6

I. 蛋… II. ①利…②张… III. 蛋白质-研究 IV. Q51

中国版本图书馆 CIP 数据核字 (2004) 第 098071 号

责任编辑:莫结胜 丁顺华 卢庆陶 / 责任校对:陈丽珠
责任印制:钱玉芬 / 封面设计:王 浩 陈 敬

科学出版社 出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

双青印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2005年1月第一版 开本:B5(720×1000)

2005年1月第一次印刷 印张:8 1/4

印数:1—3 000 字数:154 000

定价:30.00 元

(如有印装质量问题,我社负责调换〈环伟〉)

译者的话

在后基因组时代，生物学家正在对基因组的功能进行研究。蛋白质组学是生物技术的最新领域，它将对基因组功能的研究做出巨大贡献。基因只含有制造蛋白质的指令，而细胞的各种各样的生理功能是由蛋白质来完成的。蛋白质组学将组织或细胞中的蛋白质作为一个系统来研究，而不像在蛋白质化学中那样只研究蛋白质的单一组分。在不同的细胞中会有不同的蛋白质的表达，在相同细胞的不同状况中（如处在疾病状态的细胞中），也会有某些不同的蛋白质的表达。细胞只有一个基因组，但可能有许多不同的蛋白质组。蛋白质组比基因组更复杂。这种复杂性还表现在蛋白质有很多难以预测的翻译后修饰、蛋白质-蛋白质相互作用以及折叠成各种形状的三维结构。从蛋白质的氨基酸序列不一定能推测蛋白质的功能，而蛋白质组学的研究可以揭示蛋白质的表达和功能。各国科学家目前正在用蛋白质组学对人类全部蛋白质进行分类，研究人类蛋白质的相互作用，以便开发更有效的药物。

蛋白质组学面临的挑战是必须研究复杂的蛋白质体系。这要求我们分析各种各样的蛋白质，这些蛋白质大部分以修饰的形式存在并且是低丰度的。《蛋白质组学导论——生物学的新工具》一书描述了应对这个巨大挑战的工具和方法。本书介绍了分析蛋白质和肽的各种方法，重点阐述了不同质谱仪及相关数据库检索算法的基本原理和使用方法，详细描述了质谱在蛋白质组学中的应用。本书作者 D. C. 利布莱尔 (Daniel C. Liebler) 教授是有丰富蛋白质组学研究和教学经验的科学家，尤其在使用质谱技术及相关算法鉴定蛋白质方面有很高造诣。他特别重视蛋白质组学的应用，发表了一系列用质谱技术研究蛋白质修饰的文章。作者力图使本书成为一本可供有生物化学背景的学生和科学家使用的入门教材，书写流畅，深入浅出。蛋白质组学领域最权威的科学家之一 J. R. 耶茨 (John R. Yates) 教授称本书是蛋白质组学的极好的导论和综述。

本书共分三部分。第 I 部分用两章描述了蛋白质组学和蛋白质组的定义，阐述了蛋白质组学诞生和发展的基础以及在新生物学中的地位，讨论了蛋白质组与基因组的关系。第 II 部分介绍了蛋白质组学的工具和方法。在第 4 章和第 5 章讨论了蛋白质和肽的分离方法以及蛋白质的消化技术。蛋白质和肽的分离包括使用二维 SDS 聚丙烯酰胺凝胶电泳 (2D-SDS-PAGE)、制备等电聚焦、HPLC、串联液相层析和毛细管电泳。第 6 章详细描述了 MALDI-TOF 质谱仪和 ESI 串联质谱仪的结构和工作原理以及它们的优缺点，讨论了在蛋白质组学研究中如何选用不同的质谱仪。第 7 章的主题是用肽质量指纹谱鉴定蛋白质，讨论通过测定的

肽质量与数据库理论肽质量的比较进行蛋白质鉴定的方法。第 8 章到第 10 章主要阐述如何用串联质谱分析肽序列, 描述了从串联质谱谱图鉴定蛋白质的软件工具 (主要介绍 Sequest), 也介绍了如何使用 SALSA 算法采集串联质谱数据特征。第 III 部分详细介绍了质谱和相关技术在蛋白质组学中的应用, 描述了在蛋白质采集和在蛋白质表达谱的研究中, 2D-SDS-PAGE 和 MALDI-TOF 质谱以及肽的多维层析和 LC-串联质谱分析的应用 (第 11 章、第 12 章); 讨论了在鉴定蛋白质-蛋白质相互作用和蛋白质复合物以及鉴定蛋白质修饰中, 质谱以及 Sequest 和 SALSA 算法的应用 (第 13 章、第 14 章)。最后一章指出了蛋白质组学的新的发展方向, 包括新质谱仪、自动化和蛋白质微阵等。

相信本书将为从事生命科学和医学研究的专业人员, 以及学习生命科学和生物技术的本科生和研究生助一臂之力。

张继仁

序

自 1958 年克劳斯·比曼 (Klaus Biemann) 教授首先用质谱仪分析氨基酸以来, 质谱技术已有了长足发展。Biemann 最初的实验所面临的棘手问题是怎样将非极性分子引入质谱仪产生离子。1958 年以后出现的几种新型电离技术和样品导入方法, 促进了生物分子的分析, 如化学电离、电场解吸、场致电离、等离子体解吸以及快原子轰击 (FAB) 等新型电离技术, 鉴定肽和蛋白质的方法也得以发展。1987 年由于在生物分子中引入了基质辅助激光解吸电离 (MALDI) 以及电喷雾电离 (ESI), 质谱技术有了跃进。这两种电离技术也给肽和蛋白质分析带来极大飞跃, 其中一个关键质谱技术是串联质谱。

在 20 世纪 80 年代早期唐纳德·亨特 (Donald Hunt) 教授开始在肽和蛋白质序列分析中发展和应用串联质谱。FAB 是一项软电离技术, 可产生完整的质子化分子, 使得用于肽序列分析的方法得以改进。FAB 是肽序列测定的主要突破, 该技术可以使肽稳定电离, 无需通过其他方法增加肽的挥发性。FAB 与串联质谱的联用, 形成了快速肽序列测定方法学。处理复杂肽混合物时, 大多数方法采用离线 HPLC 进行分离。人们通过这种方法对许多蛋白质进行了序列测定, 并发展了许多重要的方法。然而分离方法与 FAB 的在线结合一直未能发展出可靠易行的方法。直到电喷雾电离使分离技术与质谱仪直接联用, 这个问题才得到解决。分析灵敏度的增加以及样品处理的简化和自动化使肽和蛋白质分析的各个方面都得以提高。

质谱的这些进展与全球协同进行的人类基因组序列测定很好地衔接在一起。基因组序列测定工作包括人类基因组及许多模式生物的基因组, 并已产生大量的序列信息。1993 年, 几个研究小组发现质谱数据可用来检索数据库, 以鉴定所研究的蛋白质。1994 年发展了用串联质谱数据检索序列数据库的方法, 使研究者能在“书的后面看到答案”。如果“书”是已得到序列分析的生物基因组, 答案基本上肯定是在书后面的部分。翻译后修饰和氨基酸序列改变等复杂问题可通过研究从基因组序列推出的蛋白质序列得以解决。

20 世纪 90 年代在生物科学中人们对质谱的兴趣和应用迅速增长, 质谱在新千年将会像 SDS-PAGE 一样普遍和重要。生物学家将依赖质谱判断其实验结果。如果生物学家需要使用质谱技术来分析实验, 那么他们怎样了解质谱艺术和蛋白质组学的方法呢? D. C. 利布莱尔 (Daniel C. Liebler) 教授的《蛋白质组学导论——生物学的新工具》一书可以指导我们了解质谱并且在蛋白质组学研究中使用质谱。这本书描述了通常使用的质谱仪和基本的电离技术, 这对于确定特定研

究中如何选用质谱仪的类型是重要的。对于非专业研究人员来说，使用质谱数据检索数据库是重要的改进，这样就不再需要掌握解释质谱图的技巧。本书描述了对基础检索算法的基本理解，并阐述了其局限性，最后描述了质谱在蛋白质组学中的应用。本书为研究生和所有对迅速发展的蛋白质组学基础知识感兴趣的生物学家提供了极好的蛋白质组学导论和综述。

J. R. 耶茨 (John R. Yates)
Scripps Research Institute
La Jolla, CA

前 言

本书是蛋白质组学这个新领域的导论，侧重描述怎样分析研究蛋白质和蛋白质组。尽管人们对蛋白质组学的兴趣日益浓厚，但是对蛋白质组学的工具和技术了解还很少。本书注重向生物学家介绍新工具和新方法，对生物学学生和有经验的生物学家都适用。任何学过研究生生物化学课程的人都可以很容易地理解什么是蛋白质组学以及如何研究蛋白质组。有经验的生物学家会发现本书大部分内容是熟悉的，但是这些内容被重新整合并围绕蛋白质组的研究展开阐述。

基因组序列测定、分析仪器、计算能力和易于使用的软件工具等方面的重大进展已不可逆地改变了生物学的发展方向。过去我们一直研究生物系统的单个组分，而现在可以综合地并且在精确的分子细节上研究生物系统本身。我们面对的任务是有效地利用新技术和处理大量的数据，更重要的是我们需要调整思想去理解与单一组分相对的复杂体系。

《蛋白质组学导论——生物学的新工具》这本书最早是用质谱进行肽序列分析的短期课程讲义，这门课程是由 Donald F. Hunt 博士 1998 年在北卡罗来纳州 Durham 的生物医学资源设施协会会议上讲授的。那时我的同事 Tom McClure 博士和我在亚利桑那大学毒理中心和亚利桑那癌症中心建立了一个新的蛋白质组学研究机构。Tom 参加了 Hunt 的课程。他回来后，将授课内容传授给我们几个。我们于 1998 年 8 月在亚利桑那大学开设了一个为期四天的关于质谱和蛋白质组学的训练班，共有 50 个学员接受了培训，学员包括研究生、实验室工作人员和教授。对这个训练班的热烈反响反映了以某些易接受方式将蛋白质组学的新技术和在研究中的潜在应用介绍给科学家的需要。这个训练班促进了本书的诞生。

本书是写给初学者的。我的目的是让他们熟悉蛋白质组学的重要工具和应用，所以对某些仪器和应用的描述并不是非常严谨的。本书不是实验室手册或最新技术汇编。有几本很好的书更详细地描述了蛋白质分析技术、质谱仪器和技术，以及这些技术的应用。在这个领域中研究方法的发展和应用是非常迅速的，没有哪本书是真正时新的。在我将蛋白质组学介绍给同事后令我兴奋的是同事们创造性地运用这些新技术，这将促进蛋白质组学的发展。

本书分成三部分。第Ⅰ部分介绍蛋白质组学主题，描述它在新生物学中的位置，分析蛋白质组的性质。第Ⅱ部分介绍蛋白质组学研究的工具，解释它们怎样工作。第Ⅲ部分解释这些工具怎样用来解决不同类型的生物学问题。

感谢 Jeanne Burr、Laura Tiscareno、Julie Jones、Dan Mason、Beau Hansen、Hamid Badghisi、Linda Manza、Richard Vaillancourt、Tom McClure、Arpad Somogyi 和 George Tsaprailis，他们提出了很好的建议，阅读书中各章的草稿并给出评语，提供某些图解的样品数据。感谢 Elizabeth Hedger 杰出的秘书工作。最后，感谢我的妻子 Karen 和儿子 Andrew 对我写作的支持。

D. C. 利布莱尔 (Daniel C. Liebler), PhD

目 录

译者的话

序

前言

I 蛋白质组学和蛋白质组	1
1 蛋白质组学和新生物学	3
2 蛋白质组	9
II 蛋白质组学的工具	17
3 分析蛋白质组学概述	19
4 蛋白质和肽的分析分离	21
5 蛋白质消化技术	33
6 分析蛋白质和肽的质谱仪	37
7 用肽质量指纹谱鉴定蛋白质	51
8 用串联质谱分析肽序列	57
9 用串联质谱数据进行蛋白质鉴定	63
10 SALSA:一种采集串联 MS 数据特征的算法	69
III 蛋白质组学的应用	77
11 采集蛋白质组	79
12 蛋白质表达谱	86
13 鉴定蛋白质-蛋白质相互作用和蛋白质复合物	95
14 蛋白质修饰谱	105
15 蛋白质组学的新方向	115
索引	121

I 蛋白质组学和蛋白质组

1 蛋白质组学和新生物学

1.1 新生物学

蛋白质组学是研究与基因组相对应的蛋白质组的学科。术语“蛋白质组学”(proteomics)和“蛋白质组”(proteome)是 Marc Wilkins 及其同事在 20 世纪 90 年代早期提出的,对应于描述生物中全部基因的术语“基因组学”(genomics)和“基因组”(genome)。这些“-omics”术语代表了对怎样思考生物学和生物体系工作方式的重新定义(图 1.1)。直到 20 世纪 90 年代中期,生物化学家、分子生物学家和细胞生物学家还在研究单独的基因和蛋白质或与生物化学途径相关的少量组分。那时可用的技术有 Northern 印迹法(用于基因表达分析)和 Western 印迹法(用于蛋白质分析),利用这些技术研究和分析多基因或多蛋白质是非常困难的。

三项进展形成了新生物学的基础,改变了生物学研究前景。第一项是 20 世纪 90 年代基因、表达序列标签(EST)和蛋白质序列数据库的发展。作为许多生物基因的部分信息库,这些资源很有价值。20 世纪 90 年代后期的基因组序列测定工作,阐明了细菌、酵母、线虫和果蝇的完整基因组序列,最近阐明了人类基因组完整序列。植物和其他广泛研究的动物基因组的序列最近也已完成或接近完成。这些基因组数据库是我们最终从中获取对生物体系理解的信息库。

第二项重要进展是引入易于操作的、基于浏览器的生物信息学工具。利用这些工具从上述数据库中获取信息。现在可以在几秒钟内从完整的基因组内检索特定的核酸或蛋白质序列。这样的数据库检索工具与其他工具和数据库结合利用,根据已存在的特定功能结构域和基序可以预测蛋白质产物的功能。这样一批基于互联网的免费工具使生物学家可以通过台式电脑检测基因和基因产物的结构与功能,探索大量感兴趣的生物化学问题。

第三项重要进展是寡核苷酸微阵。微阵含有在玻片上或芯片上的一系列基因专一性寡核苷酸或 cDNA 序列。将感兴趣样品的 DNA 混合物荧光标记后与微阵进行杂交,可以一次探测几千个基因的表达。一个微阵可以代替几千个 Northern 印

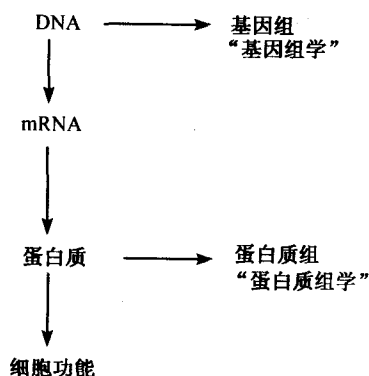


图 1.1 基因组学和蛋白质组学的生物化学关系

迹法分析，可以在做一次 Northern 印迹的时间内完成。通过使用双色荧光探针标记，两个不同样品的基因表达谱可直接在一个玻片或芯片上进行比较。

图 1.2 是一块含有酿酒酵母基因组中 6 000 个基因单一序列的玻片。这样的微阵可以测定酵母基因组所有基因的表达。显然，这使我们面临新生物学的巨大挑战。我们可以观看整个系统，但这几千个数据点所包含的信息超出了我们直观解释的能力。新的组合算法、自组作图和类似的工具等最新的方法有助于生物学家理解这些数据。

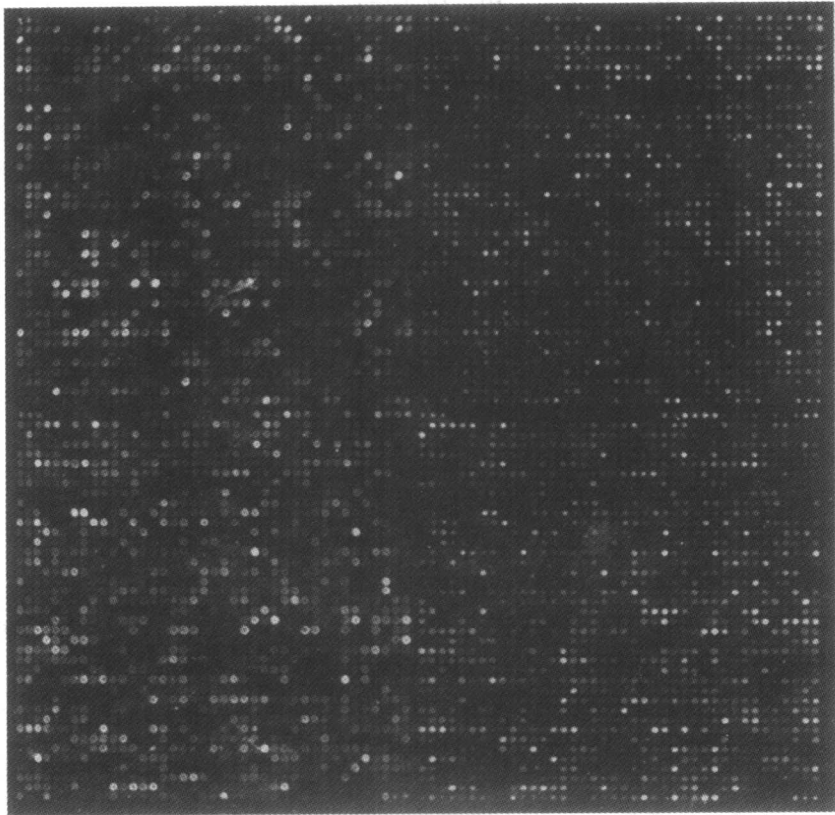


图 1.2 酵母基因组芯片

该酵母 cDNA 微阵由斯坦福大学 Patrick Brown 博士的实验室制作 (<http://cmgm.stanford.edu/pbrown/>)。

微阵带来的最重要的变化是使生物学家进行“宏观”思考。细胞有成千上万的以不同组合方式进行的基因表达。细胞的生死是由这些基因的表达和其蛋白质产物的活性决定的。无论是跨膜受体、转录因子、蛋白激酶或是伴侣分子，每种蛋白质所表达的功能只有在同一细胞内其他蛋白质的功能和活性同时表达时才有意义。因此生物学家正在努力做宏观思考，去理解系统而不只是理解组分，寻求复杂性的意义。

1.2 蛋白质组学？它不过是我们过去称之为蛋白质化学的学科！

对新思想、术语和方法人们通常说它们其实不是新发展来的，因此解释蛋白质组学与蛋白质化学的不同是很重要的。表 1.1 对各自的主要特征作了小结。蛋白质化学包括研究蛋白质的结构和功能，通常涉及物理生物化学或机械酶学。研究工作通常包括完整序列测定、结构测定以及进行结构控制功能的模型研究。物理生物化学家和酶学家在同一时间内只研究一个蛋白质或多亚基蛋白质复合物。

表 1.1 蛋白质化学和蛋白质组学的不同

蛋白质化学	蛋白质组学
单一蛋白质	复杂混合物
全序列分析	部分序列分析
强调结构与功能	强调通过数据库匹配进行蛋白质鉴定
结构生物学	系统生物学

蛋白质组学研究多蛋白质系统，重点研究作为一个大系统或部分网络的组成的多个不同蛋白质的相互作用。蛋白质组学需进行复杂混合物的分析，不是通过完整序列测定进行鉴定，而是在数据库匹配工具帮助下进行部分序列测定。蛋白质组学的内容是系统生物学，而不是结构生物学。换句话说，蛋白质组学的要点是鉴定系统的行为而不是任何单一组分的行为。

1.3 我们能测定基因表达，为什么还要有蛋白质组学？

基因微阵提供了细胞中大量或全部基因表达的快速检测。然而从 mRNA 水平并不一定能预测细胞中相应蛋白质的水平。各种 mRNA 不同的稳定性和不同的翻译效率能够影响新蛋白质的产生。蛋白质形成后，在稳定性和转换速度上有很大不同。许多参与信号传导、转录因子调节和细胞周期控制的蛋白质迅速转换，这是其活性调节的一种方式。mRNA 水平没有告诉我们相应蛋白质的调节状态，蛋白质的活性和功能常有一些内源翻译后的改变，也会因环境因素而改变。

1.4 蛋白质组学：对分析的挑战

如何同时测定一个生物中大量或全部基因的表达似乎已通过引入 cDNA 或寡核苷酸微阵得以解决。用 DNA 微阵和相关方法分析基因表达依赖于两个重要工具：聚合酶链反应（PCR）和寡核苷酸与互补序列的杂交。但是没有类似的工具用于蛋白质分析。首先，没有蛋白质 PCR 等价物。目前不可能有多肽分子以类似于核苷酸通过 PCR 复制的方式复制。少量的寡核苷酸可以通过 PCR 进行扩增，而少量的蛋白质必须在没有任何扩增的情况下进行测定和分析。

第二，蛋白质不能专一性与互补氨基酸序列杂交。Watson-Crick 碱基配对

允许寡核苷酸与互补序列杂交。一个特定的互补寡核苷酸序列可以作为高度专一性探针，一个特定的 mRNA 或其他核酸片段可以与之结合。这种专一性允许在微阵上有一个特定的点以便识别单一序列。尽管抗体和寡核苷酸结合子 (aptamer, 也称适体) 可以识别特定的肽或蛋白质，但是这种识别不能简单地根据序列来预测，而寡核苷酸的杂交则可以根据序列来预测。

另一个蛋白质组学的特有问题是细胞中每一个蛋白质产物并不一定只有一种分子实体。这是由于蛋白质有翻译后修饰。修饰的内容和变化随不同的蛋白质、细胞的调节机制和环境因子而变化，许多蛋白质以多种形式存在。对任何特定基因的多种蛋白质产物进行检测和区分的必要性使蛋白质组学在分析方面更具挑战性。

蛋白质组的分析需要一套不同于基因表达分析的工具，能够对修饰和非修饰的蛋白质进行检测和定量的分析。我们怎样应对这项任务是这本书的主题。

1.5 蛋白质组学的工具

尽管上面描述了分析蛋白质组学的不利条件，但是鉴定蛋白质组及其组分实际上可以完成。这是由于以下 4 种重要工具的发展和结合使用给研究人员提供了灵敏性和专一性较高的识别和鉴定蛋白质的方法。

第一种工具是数据库。蛋白质、EST 和基因组序列数据库共同提供了生物表达全部蛋白质的完整数据库目录。例如，根据对果蝇的所有编码序列的分析，我们知道有 110 个果蝇基因编码具有 EGF 类结构域的蛋白质，87 个基因编码具有酪氨酸激酶催化结构域的蛋白质。在进行果蝇蛋白质组学研究时，我们检索大量已知的可能蛋白质结构域。当用限定的序列信息，甚至原始质谱数据（见下文）进行检索时，我们可以根据质谱数据与数据库的匹配情况鉴定蛋白质组分。

第二种工具是质谱 (MS)。质谱仪的使用在过去十年中有了极大的革新，在发展为分析生物分子，特别是分析蛋白质和肽的高灵敏度和高可靠性上达到顶点。MS 仪器的使用可提供三类分析，这三类分析在蛋白质组学分析中都非常有用。首先，MS 可以进行 100 kDa 或更大完整蛋白质的精确质量测定。估计蛋白质质量的最好方法是 MS 分析，而不是测定蛋白质在十二烷基硫酸钠-聚丙烯酰胺凝胶电泳 (SDS-PAGE) 的迁移。高精度蛋白质质量测定的应用有限，因为它们往往不够灵敏。净质量对精确鉴定蛋白质往往是不充分的。其次，MS 也能对蛋白质水解消化产生的肽进行精确的质量测定。相对于完整蛋白质质量测定，肽质量测定可以有高灵敏度和高质量准确度。可以直接用肽质量测定数据在数据库中进行检索，这样常常可以确切鉴定靶蛋白质。最后，MS 可以对蛋白质水解消化得到的肽序列进行分析。目前认为 MS 是肽序列分析中的最新技术。MS 序列数据为蛋白质鉴定提供了最有力和最精确的方法。

蛋白质组学的第三个必要工具是对数据库中特定蛋白质序列与 MS 数据进行比对的各种软件。前面提到从 MS 数据可以测定序列，但是这种从头开始分析成

百上千的谱图时是一项费时费力的任务。蛋白质组学软件将未分析的 MS 数据,在特定算法的帮助下与蛋白质、EST 和基因组序列数据库的序列相比对,自动检测大量用于蛋白质序列匹配的 MS 数据。然后研究人员检查自动检测的结果,估计数据的质量,所用的时间比手工解释每一张谱图要少得多。

蛋白质组学的第四种必需工具是蛋白质的分析分离技术。在蛋白质组学中蛋白质分离有两个目的。第一,通过将蛋白质混合物分离成单一蛋白质或蛋白质小组以简化复杂蛋白质混合物。第二,蛋白质的分离分析可以比较两个样品蛋白质的不同表现,研究者可以标记用于分析的特定蛋白质。2D-SDS-PAGE 是最广泛用于蛋白质组学的技术。2D 凝胶电泳也许是在复杂样品中分离蛋白质的最好单项技术。其他的蛋白质分离技术,包括 1D-SDS-PAGE、高效液相层析 (HPLC)、毛细管电泳 (CE)、等电聚焦 (IEF) 和亲和层析,也都是分析蛋白质组学的有用工具。最有力的技术是将不同的蛋白质和肽分离技术结合为多维技术。例如,离子交换液相层析 (LC) 与反相 (RP)-HPLC 的串联是分离复杂肽混合物的有力工具。

这四种工具的结合形成了蛋白质组学当前的技术,每一种工具在技术上都发展迅速。在本书的后面几章我们将讨论每一种分析工具。

1.6 蛋白质组学的应用

蛋白质组学技术确实很新颖,但是鉴定蛋白质组究竟是为了什么呢?根据目前的实践,蛋白质组学包括 4 项主要应用,它们是:①采集;②蛋白质表达谱;③蛋白质网络谱;④蛋白质修饰谱。下面对上述每一项进行简短定义,在本书其后各章将详细讨论。

采集是鉴定样品中所有(或尽可能多)的蛋白质。采集主要是直接对蛋白质组进行分类,而不是通过基因表达(如通过微阵)数据来推断蛋白质组的组成。蛋白质组学中采集需要耗费大量的劳动,以使蛋白质得到最大程度的分离,然后使用 MS 和相关的数据库以及软件工具进行鉴定。有几种采集方法,每一种都有其优点。这些方法的联用可直接分析证实那些只能从基因表达数据推断的数据。

蛋白质表达谱是鉴定生物或细胞特定状态(如分化、发育状态或疾病状态)下蛋白质的表达或药物、化学或物理刺激下蛋白质的表达。表达谱其实是特殊的采集形式,分析中比较一个特定系统的两种不同状态。例如,比较正常细胞和病理细胞中哪些蛋白质有不同的表达。这种信息对于检测药物治疗的潜在靶子极为有用。

蛋白质网络谱是在生物系统中测定蛋白质之间相互作用的蛋白质组学方法。大多数蛋白质在执行功能时与其他蛋白质密切相关。这些相互作用决定蛋白质功能网络(如信号传导级联过程和复杂的生物合成或降解途径)的功能。大多数蛋白质-蛋白质相互作用是通过体外纯化的蛋白质和用酵母双杂交系统获得。通过亲和俘获配对技术与分析蛋白质组学方法相结合,蛋白质组学可以鉴定更复杂的