



生命科学前沿及应用生物技术

工业蛋白质组学

——在生物技术和制药中的应用

(美) D. 菲格斯 主编

钱小红 贺福初 主译

Industrial Proteomics

Applications for Biotechnology and Pharmaceuticals



科学出版社

生命科学前沿及应用生物技术大系·典藏版

工业蛋白质组学

——在生物技术和制药中的应用

Industrial Proteomics

Applications for Biotechnology and Pharmaceuticals

〔美〕D. 菲格斯 主编

钱小红 贺福初 主译

科学出版社

北京

图字：01-2005-6306 号

内 容 简 介

应用生物技术大系和现代生命科学前沿系列图书分别被列为“十一五”和“十二五”国家重点图书出版规划项目。本丛书针对生命科学领域前沿重点发展方向以及应用生物技术领域的新成果、新思路、新方法和新技术，全面展示了其最新的发展动态，涵盖了基础理论和主要技术方法，呈现了新的概念与理论、技术，在更深层次上阐明了生命的本质规律，给人们提供了新的认识生命本质的手段，也为生物技术服务于人类开辟了新的途径。涉及领域包括生物医药、干细胞技术、工业微生物学、蛋白质及蛋白质组、系统生物学、合成生物学、生物材料、农业生物技术、环境生物技术、海洋生物技术、生物资源与安全等。

Industrial Proteomics: Applications for Biotechnology and Pharmaceuticals

Copyright ©2005 by John Wiley & Sons, Inc. All rights reserved.

Authorized translation from the English language edition by John Wiley & Sons, Inc.

图书在版编目（CIP）数据

生命科学前沿及应用生物技术大系：典藏版/舒红兵等编著. —北京：科学出版社, 2016

ISBN 978-7-03-047487-2

I. ①现… II. ①舒… III. ①生命科学—研究②生物工程—研究 IV. ①Q1-0②Q819

中国版本图书馆 CIP 数据核字(2016)第 043876 号

责任编辑：王 静 李 悦

责任印制：张 伟 / 封面设计：刘新新

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>

北京厚诚则铭印刷科技有限公司 印刷

科学出版社发行 各地新华书店经销

*

2016 年 7 月第 一 版 开本：787×1092 1/16

2016 年 7 月第一次印刷 印张：2108

字数：49 985 000

定价：8900.00 元

(如有印装质量问题，我社负责调换)

译者的话

21 世纪是生物经济的时代,生物产业将成为高新技术产业的龙头。蛋白质组研究为高通量挖掘重要功能蛋白质、批量发现药靶及原创性新药的研究与开发提供了新的思路 and 手段。

本书全面、详细地介绍了蛋白质组学在药物设计、蛋白质药物标靶的规模化识别与鉴定、生物标志物的通量筛选等诸多领域的研究与应用,填补了国内同类出版物的空白。

本书的译者主要是多年从事蛋白质组研究的专家、学者,具有丰富的一线研究经验。我们翻译和出版此书,希望能为我国从事蛋白质组学研究的相关人员提供一定的理论指导及技术参考。

书中涉及的技术层次非常新,应用领域又非常广,给翻译工作带来了很大的挑战,译文中难免会有疏漏,衷心希望得到各位同道、读者的指正。

序 言

在 John 的《尚需探索》一书中,《自然》杂志前主编 John-Maddox 先生阐述了深入了解生物分子多样性的必要性并总结说:“在整个自然界中,现象多样性掩盖着其本身简单的本质。”基因组学技术的发展,尤其是这一技术能够迅速准确地获得生物信息,是解密人体生物学的第一步。但是,自然界特别擅长采用复杂的现象来掩盖其简单的本质,因此基因组学的技术手段显得有些力不从心。从分子水平上理解人类和其他生物体复杂多样的生物学过程给我们的科学与技术研究提出了挑战,这是单单一个基因组测序所不能够达到的。研究细胞内复杂多样的生物分子以及迅速发展的一些特异通路和生物分子需要国际上众多研究者的共同协作。

蛋白质组通常被当作基因组在蛋白质水平上的一个补充来介绍,但是,它们的相似之处仅限于此。由于自身和外界压力的作用,蛋白质进行着包括产生、修饰、翻译和降解等复杂的加工过程,蛋白质组自始至终进行着变化。可能听起来这一过程显得混乱复杂,实际上它是蛋白质在不断的调控下而形成的一个组织有序而高效的体系。

长期以来,蛋白质的研究多是对特定时间的单一蛋白质进行研究。经典的生物化学和分子生物学方法一直引导着我们对细胞内各种过程的理解,同时也给我们提供了大部分的药物靶标。蛋白质组学被界定为研究蛋白质组,通常认为需要一些高通量的方法。虽然我们探索蛋白质组学的方法相比较基因组学来说更局限,但是过去的十几年来,蛋白质组学的研究还是取得了巨大的成就。

许多生物技术公司和大部分的医药公司迅速地接受并采用蛋白质组学的方法来弥补基因组学研究的不足。蛋白质组学在寻求靶标、发现药物和临床诊断技术方面具有举足轻重的作用。最初的工作大多是采用二维电泳与质谱鉴定相结合的方法寻找靶标。OGS 和大规模生物学是以二维电泳为基础的蛋白质组学在工业上应用的先锋。由于结合了分子生物学技术、噬菌体展示技术以及质谱技术,以二维电泳技术为核心的蛋白质组学迅速扩展到表达蛋白质组学、功能蛋白质组学和结构蛋白质组学,蛋白质组学迅速的发展吸引了越来越多的公司加入到这一竞技场中。

虽然已有市售的一些关于蛋白质组学的专著,但是我们认为还需要一本专注于蛋白质组学在工业中应用的书。虽然在技术方法上可能非常相似,但是蛋白质组学在工业上应用的范围和方向与科学研究迥然不同。蛋白质组学在工业中的应用主要专注于对于靶标的筛选和医药的工艺过程。因此这就要求蛋白质组学的操作过程不仅实用性强,特性优良而鲜明,而且有严格的质量控制,能够产生统计学上明显的结果。同时,蛋白质组学在操作的过程中能够处理大量的样品。例如,一个简单的临床诊断研究可能要求通过表达蛋白质组的方法来分析从最少 36 个样本到数以万计的复杂样本。

本书包含十一章,每一章都着重阐述蛋白质组学在工业应用中的某一特定领域。第一章主要介绍了以质谱技术为基础的蛋白质组学中的一些基本内容。功能蛋白质组的内容在第二和第三章阐述。在第二章中讨论了以质谱技术为基础的蛋白质相互作用谱的构

建。在第三章 Annan 和 Zappacosta 讨论了蛋白质的翻译后修饰,尤其是磷酸化修饰。结构蛋白质组学在第四和第五章中进行阐述。在第四章,来自 Syrrx 和 Active Sight 的 Tarrin, Jennings 和 McRee 论述了高通量晶体学的应用以及基于 *in silico* 方法的以结构为基础的药物设计。在第五章,来自 ExSAR 公司的 Hamuro, Weber 和 Griffin 描述了氢/氘交换质谱技术在高通量蛋白质结构研究中的应用。

蛋白质组学最初的应用在于靶标的发现。在第六章,来自 Eli Lilly 的 Hale, Ou, Shiyonov, Knierman 和 Ludwig 讨论了蛋白质组学技术在鉴定以及对蛋白质靶标的确认方面的应用。

蛋白质组学的最新应用主要是疾病或与药物相关的生物标志物的发现。在第七章中,来自 MDS 医药服务处的 Masé 和 Gibbs 对于生物标志物的发现以及确认提出总的见解。同时,来自 GeneProt 的 Rose 在第八章详细描述如何应用蛋白质组学方法发现血浆生物标志物。

蛋白质组学还可以用来研究小分子(例如,药物),尤其是寻找与药物相互作用的蛋白质。在第九章中,来自 Xencor 的 Doberstein, Hammond 和 Hubert 提出了化学基因组学与化学蛋白质组学,并且讨论了不同的技术方法。

在第十章中,来自丹麦 MDS Inc. 的 Ahren, Jespersen 和 Schandorff 提出了一个蛋白质易使用的生物信息学方法和在发展这一方法中需要注意的重要因素。在第十一章中,来自 Promab 的 Wilson 和 Nock 通过介绍不同的方法和讨论其中的挑战与成功给我们展示一个充满希望的蛋白质芯片领域。

我认为蛋白质组学未来在系统生物学的发展和医药领域的临床前和临床诊断应用中将会大有作为。本书不仅仅是一本写给工业领域同仁的书。通过使用良好特性的质量控制(QC)体系来产生统计学上有意义的结果从而使得蛋白质组学在系统的方式下实施,这一概念在科学研究与发现中同样重要和有效。蛋白质组学在工业中的经验将会在系统生物学和药物发现中有着广泛的应用。

(钱小红 译)

本书主要参与者

Christian Ahren MDS Inc. -Denmark, Staermosegaardsvej 6, Odense M, DK-5230, Denmark

Roland S. Annan Proteomics and Biological Mass Spectrometry, Department of Computational, Analytical, and Structural Science, GlaxoSmithKline, King of Prussia, Pennsylvania

Steve Doberstein Five Prime Therapeutics, Inc. ,951 Gateway Boulevard, South San Francisco, CA 94080

Daniel Figeys Department of Biochemistry, Microbiology, and Immunology, University of Ottawa, Ottawa, Canada

Bernard F. Gibbs Applied R+D Department, Early Clinical Research and Bioanalysis, MDS Pharma Services, Montreal, Quebec, Canada

Patrick R. Griffin ExSAR Corporation, 11 Deer Park Drive, Suite 103 Monmouth Junction, New Jersey, 08852

John E. Hale Enabling Biology Department, Lilly Research Labs, Indianapolis, Indiana

Philip W. Hammond Xencor, 111 W. Lemon Ave. , Monrovia, California, 91016

Yoshitomo Hamuro ExSAR Corporation, 11 Deer Park Drive, Suite 103 Monmouth Junction, New Jersey, 08852

René S. Hubert Xencor, 111 W. Lemon Ave. , Monrovia, California, 91016

Andy J. Jennings Syrrx, Inc. 10410 Science Center Dr. San Diego, California, 92121

Hans Jespersen MDS Inc. -Denmark, Staermosegaardsvej 6, Odense M, DK-5230, Denmark

Michael D. Knierman Enabling Biology Department, Lilly Research Labs, Indianapolis, Indiana

James R. Ludwig Enabling Biology Department, Lilly Research Labs, Indianapolis, Indiana

Robert Masé Applied R+D Department, Early Clinical Research and Bioanalysis, MDS Pharma Services, Montreal, Quebec, Canada

Duncan E. McRee Active Sight, 4045 Sorrento valley Blvd. , San Diego, California, 92121

Steffen Nock Promab Biotechnologies, Inc. , San Leandro, California, 94577

Weijia Ou Enabling Biology Department, Lilly Research Labs, Indianapolis, Indiana

Keith Rose GeneProt, Inc. , 2 rue Pre-de-la-Fontaine, CH-1217 Meyrin/Geneva, Switzerland

Soeren Schandorff MDS Inc. -Denmark, Staermosegaardsvej 6, Odense M, DK-5230, Denmark

Pavel Shiyarov Enabling Biology Department, Lilly Research Labs, Indianapolis, Indiana

Leslie W. Tarr ActiveSight, 4045 Sorrento valley Blvd. , San Diego, Carlifornia, 92121

Patricia C. Weber ExSAR Corporation, 11 Deer Park Drive, Suite 103 Monmouth Junction, New Jersey, 08852

David S. Wilson Promab Biotechnologies, Inc. , San Leandro, California, 94577

Francesca Zappacosta proteomics and Biological Mass Spectrometry, Department of Computational, Analytical, and Structural Science, GlaxoSmithKline, King of Prussia, Pennsylvania

目 录

译者的话

序言

本书主要参与者

第一章 蛋白质组学的基本概念	1
前言	2
经典和功能蛋白质组学中的蛋白质组操作方法	5
蛋白质纯化	5
二维电泳的蛋白质分离	11
蛋白质加工	15
蛋白质鉴定	16
蛋白质组学当前面临的挑战	38
参考文献	43
第二章 蛋白质 蛋白质相互作用绘图	51
前言	51
方法说明	52
实验设计	62
结论	65
参考文献	66
第三章 蛋白质翻译后修饰:质谱分析磷酸化位点	69
前言	69
实验	70
结果	72
结论	82
致谢	82
参考文献	83
第四章 高通量晶体学方法与计算机方法用于基于结构的药物设计	87
引言	87
高通量构建体的选择、克隆和微量筛选	89
蛋白质表达	89
高通量机器人纳量体积的结晶和成像	90
高通量 X 射线数据收集	92
结构测定和优化	94
数据追踪	95
结构分析	95

对接	97
药物设计	100
结构预测	101
组织蛋白酶 B(cathepsin B)抑制剂的设计	102
参考文献	103
第五章 氢/氘交换质谱在高通量蛋白质结构分析中的应用	107
前言	108
蛋白质结构和动力学分析中的氢/氘交换理论	108
氢/氘交换技术综述	110
氢/氘交换结合裂解分析方法的应用	119
结论	125
参考文献	126
第六章 蛋白质组学技术用于蛋白质药靶的识别与鉴定	130
引言	130
用于药靶发现的工具	131
验证	137
信息学	139
结论	141
参考文献	141
第七章 应用蛋白质组学方法发掘生物标志物	147
前言	147
基于质谱的蛋白质组学方法	149
磷酸化蛋白质组分析	152
蛋白质复合体和信号通路	153
基于亲和标记的质谱方法	155
定量蛋白质组学	160
疾病机制与疾病诊断学	164
生物标志物在新药研发中的应用	165
总结	168
参考文献	169
第八章 工业化规模人血浆的蛋白质组学分析	175
前言	175
工业化规模的定义	176
样本的选择、收集和随机混合	177
高丰度蛋白质的去除	178
低分子质量蛋白质和高分子质量蛋白质	178
基于凝胶的蛋白质分离	179
整体蛋白和肽段的分离	179
质谱	181

生物信息学.....	182
结论.....	183
致谢.....	183
参考文献.....	183
第九章 化学基因组学:靶标展示	185
简介:化学遗传学和化学基因组学	185
化学基因组学技术.....	186
讨论和结论.....	194
参考文献.....	195
第十章 蛋白质组生物信息学	198
引言.....	198
从基因组到蛋白质组——数据复杂性增加所引发的挑战.....	200
复杂混合物和质谱分析——数据.....	201
自动化和高通量.....	204
蛋白质序列数据库:全面性和冗余性	205
减少样品和数据复杂性的策略.....	207
促进大规模功能蛋白质组学分析的数据仓库整合策略.....	209
架起实验结果和生物信息学之间的桥梁.....	210
蛋白质组学数据的交换.....	213
总结和展望.....	213
致谢.....	214
参考文献.....	214
第十一章 蛋白质芯片	217
引言.....	217
芯片的配置.....	218
固相芯片与液相芯片.....	218
蛋白质固定表面的优化.....	221
点样技术.....	222
检测技术.....	223
内容物.....	224
表达谱.....	226
总结和展望.....	229
参考文献.....	230
索引	234
彩版	

注:本书原版中的一些图片不清晰,本中译版继续沿用。读者若需清晰图片,请参照图注中的出处信息,特此说明。

第一章 蛋白质组学的基本概念

Daniel Figeys

生物化学、微生物学和免疫学系

渥太华大学，渥太华，加拿大

前言	2
从 DNA 到 RNA	4
从 RNA 到蛋白质	5
经典和功能蛋白质组学中的蛋白质组操作方法	5
蛋白质纯化	5
细胞培养	7
其他来源的样品	7
细胞亚群	9
亚细胞组分	9
二维电泳的蛋白质分离	11
二维电泳分离的基本原则	11
通过二维电泳来检测蛋白质的分离	13
处理二维电泳图谱的分析软件	14
蛋白质加工	15
蛋白质的印迹消化	15
蛋白质的胶内消化	16
蛋白质鉴定	16
Edman 降解	16
通过质谱进行蛋白质的鉴定	16
杂合仪器	21
基于 MS/MS 的方法进行蛋白质的鉴定中的生物信息学方法	35
蛋白质组学当前面临的挑战	38
经典蛋白质组学	38
无凝胶分析：另一种差别显示	39
无胶分析和基于单肽鉴定的傅氏转换质谱	41
参考文献	43

缩写列表：

2D	二维电泳
β -gal	β -半乳糖苷酶

bp	碱基对
CZE	毛细管电泳
ESI	电喷雾离子化
EST	表达序列标签
FTMS	傅里叶变换质谱
HPLC	高压液相色谱
i. d.	内径
ICAT	同位素标记亲和标签
IPG	固定 pH 梯度
MALDI	基质辅助激光解吸电离离子化
MHC	主要组织相容性复合体
Micro-ESI	微喷 电喷雾离子化
MS/MS	串联质谱
MW	分子质量
Nano-ESI	纳喷 电喷雾离子化
o. d.	外径
ORF	可读框
ppm	10^{-6}
psi	磅力每平方英寸
PVP-40	聚乙烯吡咯烷酮
RF	电极射频
SAGE	基因表达的系列分析
SNP	单核苷酸多态性
TOF	飞行时间

前言

20 世纪科学的迅速发展已经给我们的日常生活造成了直接的影响。医药和生物科学领域的众多发现大大提高了工业国家人们的生活质量和预期寿命 (图 1.1)。在 20 世纪初, 医药与生物主要侧重于描述肉眼可见的现象。但到 20 世纪末, 科学家们已经接近微米及纳米级的微观世界, 并对我们自身的生物过程的意义有了深入的了解。新世纪的曙光引领着我们用新视野来理解系统生物学以及复杂的多基因病, 同时我们需要进一步的加强对海量的基因组学信息的积累。

在 20 世纪的最后 25 年, 少数的先知者意识到制造 DNA 测序仪所需的各种部件都可以得到, 因而人类基因组测序计划可以实施。期待着新技术的出现从而能大大加快测序步伐, 各种基因组测序计划开始进行了。但是, 技术的创新显然花去了比想像更多的时间。直到 20 世纪末, 新技术的出现才大大加快大规模基因组测序计划的步伐 (图 1.2)。

值得注意的是, 在大规模基因组测序中机械化操作应运而生 (Dovichi, 1997)。机

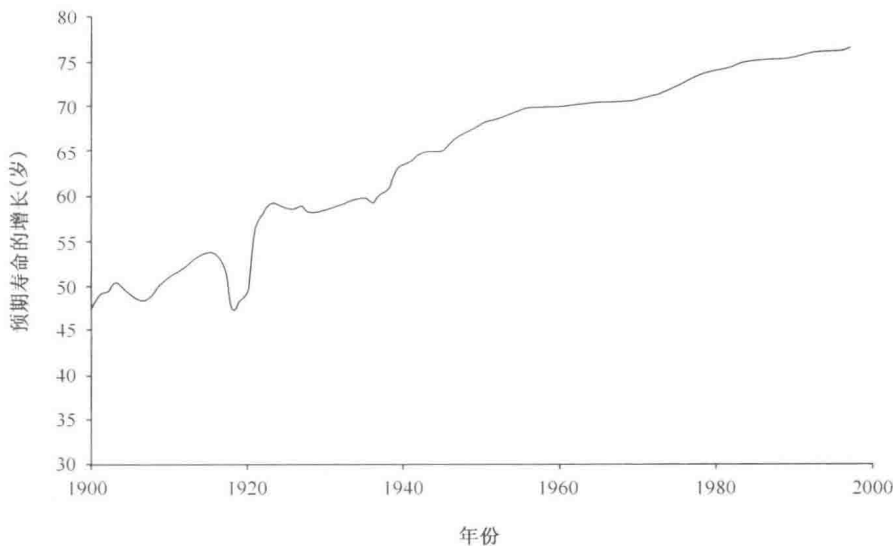


图 1.1 1900~1990 年美国预期寿命示意图

数据根据性别与种族取平均值，以 10 年为最小横坐标值。数据来源于《美国重大统计数据报告》，47 卷，28 期，1999 年，12 月 13 日出版，美国健康与人口部

械化操作对于我们发展复杂的方法来增强我们对疾病与生物的了解提供了基础。此外，科学家们很快的意识到复杂的生物学过程并不能仅仅通过简单的基因组测序来解决。很明显，目前尚待研究的一些疾病都非常的复杂，他们通常是由一系列的基因所引发的，他们的发病是由于个体的遗传学缺陷所造成的。基因组测序能够迅速的提供基因组各个部分的测序结果，但是却不能够有效地将这些部分进行合理的组合。这种不足造成了大量测序信息的获得与相对滞后的信息处理工具之间的极大不平衡，这些信息处理工具要对生物过程中的生物分子进行大规模的研究。

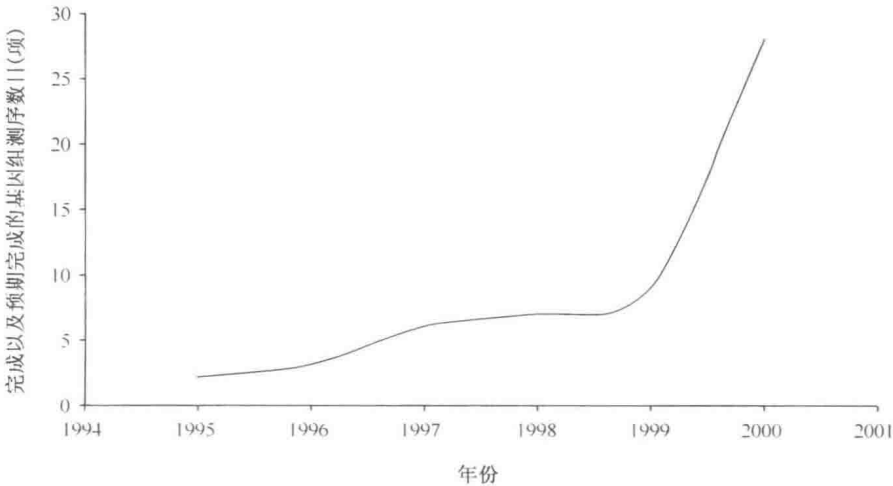


图 1.2 1995~2004 年已经完成以及预期完成的基因组测序计划

如图中所示，自 1999 年以来，基因组测序工作极其明显地开始增多

传统的生物化学方法基于对翻译后产物的功能与相互作用进行了解。获得人类基因组的价值关键在于能够找到相关的方法将基因组学的信息联系到药物发掘和诊断中。发展高通量的方法用于评估转录及翻译产物的表达水平,使细胞内加工过程的研究更加系统将会帮助我们将基因组学信息与药物发现及诊断之间联系起来。

从 DNA 到 RNA

目前有一系列的相关方法可以用来迅速准确地定量分析 RNA 的表达水平。例如, DNA/RNA 阵列以及基因表达的系列分析 (SAGE) 的出现给全部或者部分分析 RNA 的表达水平提供了方便 (Desprez et al., 1998; Marshall and Hodgson, 1998; Ruan et al., 1998; Service, 1998; Velculescu et al., 1995; Madden et al., 1997; Matsumura et al., 1999; Neilson et al., 2000; Lal et al., 1999; Stein et al., 2004; Weeraratna et al., 2004)。利用这些技术可以获得不同的 RNA 表达谱,比较不同的细胞或者不同的细胞状态时,基因表达的上调或者下调就能够检测到。RNA 与基因组的序列仅一步之遥,但通过研究 RNA 已经获得大量的信息。

很显然,这些高通量的 DNA/RNA 筛选技术给那些不同表达的基因发掘提供了一个快速而定量的方法。但是,事实上仅仅是 RNA 的表达不足以有效的了解一些生物学过程以及基因的功能。从不同的研究小组 (Gygi et al., 1999b; Gygi and Aebersold, 1999; Futcher et al., 1999) 中得到的一些数据表明, RNA 的表达与蛋白质水平上的变化有着很差的线性关系 (图 1.3),但也有其他的一些实验又表明二者有着很好的关联性 (Kern et al., 2003)。这是因为在蛋白质表达及表达后的修饰方面存在着众多的调控作用所造成的。

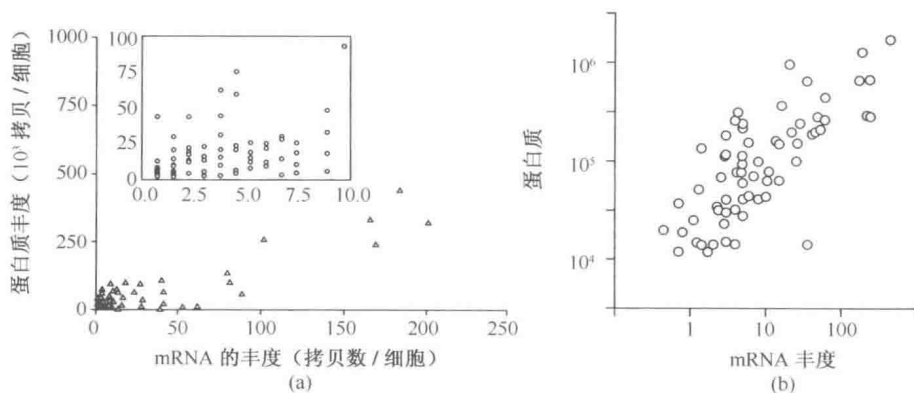


图 1.3 啤酒酵母的 mRNA 与蛋白质水平的相关性。

(a) 酵母中 106 个基因的蛋白质与 mRNA 水平的相关性。小图显示大图中低参数部分。Pearson 产物参数相关性显示全部的数据为 0.935, 小图的数据为 0.356。经作者同意, 引自 Gygi et al., (1999)。(b) 酵母中蛋白质丰度与校对过的 mRNA 丰度的相关性分析。Pearson 产物参数相关性为 0.76。经作者同意, 引自 Futcher et al., (1999)

而且,最近的一项关于酵母基因组的大规模研究清楚地显示 DNA 芯片结果与蛋白

质表达水平之间比较差的相关性 (Ross Macdonald et al., 1999)。这项研究通过采用转座子标记及基因破坏发现了 31 个减数分类基因, 这 31 个基因的蛋白质表达水平的检测通过其序列内融合入 lacZ 而检测 β -半乳糖苷酶活性。在这 31 个减数分裂基因中, 只有 17 个基因已经有报道在纺锤体形成的过程中有被诱导从而有 2 倍以上的表达。通过采用包含酵母中所有已报道的开放阅读框的 DNA 微阵列分析也证实了这一结果。但是对于其他的减数分裂基因来说, DNA 微阵列分析没能够显示出在减数分裂过程中这些基因有任何的被诱导表达的倾向。RNA 的表达水平与蛋白质的表达水平之间呈现一个复杂的关系, 因此仅靠 RNA 的表达状态来预测细胞学功能是不正确的。可以肯定的是, 通过一些下游的调控机制, 一个基因的蛋白质表达水平在 RNA 表达水平没有任何变化的情况下能够发生非常明显的变化。

从 RNA 到蛋白质

虽然科学家们已经在发展有效的基因组学技术方面做了大量的工作, 但是在生物学研究的探索中, 对于蛋白质的研究是不可避免的。研究蛋白质的原因不仅仅在于 RNA 与蛋白质的表达水平缺乏相关性。转录后的修饰, 蛋白质的翻译后剪切, 蛋白间的相互作用只是阐明蛋白质水平上复杂性的几个方面而已。

蛋白质的研究一直局限于小量的操作, 部分的原因在于缺乏明确而简单的鉴别特定蛋白质的方法。我们在进行实验的时候, 必须小心的操作从而使只有我们感兴趣的蛋白质被分离出来。在过去的 10 年由于技术的发展使得对蛋白质进行大规模分析和鉴定得以实现并得到了很大的改善 (Issaq et al., 2002; Wang and Hanash, 2003)。已经取得的成绩为进一步深入研究基因组相关蛋白质 (蛋白质组) 开启了大门。

经典和功能蛋白质组学中的蛋白质组操作方法

二维电泳最常用于蛋白质表达谱研究中, 它多用来分析在不同的条件下的不同的蛋白质表达。对蛋白质组合理保存从而获得一个真正有代表性的二维电泳结果是非常重要的, 也是蛋白质组学研究中最重要的一步。例如, 从环境中分离细胞和裂解细胞的过程中必须要足够的小心从而减少样品提取操作过程中对于蛋白质组状态的影响。在处理一个蛋白质组的过程中经常发生一些小的错误, 从而严重影响实验的结果。因此, 样品的获得过程是评价样品可靠性的先决条件。

蛋白质纯化

从细胞裂解液中提取蛋白质是建立稳定的蛋白质组的一个重要步骤。众所周知, 一旦细胞被裂解后, 一些蛋白酶就会从某些细胞器中释放出来, 从而与各种蛋白质相互作用, 使得蛋白质组被迅速的降解。幸运的是, 有许多好的方法可以使从细胞中提取的大多数蛋白质得到保护。对于可溶性蛋白质的获得可以简单地将细胞裂解之后收集上清。不同的细胞裂解的方法可以从 www.expasy.ch 网站上获取。最好选择能够与固定 pH

梯度的胶条相配的最简单的提取方法（例如最小的盐浓度和离子表面活性剂成分）。

目前的蛋白质提取方法对于所有的生物样品来说还不是通用的。考虑到提取的蛋白质组中的蛋白质代表性，它们都有其局限性。首先，疏水性蛋白质不容易被提取并用二维电泳来进行分离（Wilkins et al., 1998）。图 1.4 所示二维电泳上可见的啤酒酵母的蛋白质数目相对于 gravity 系数的疏水性范围的关系。很明显，自然界可以预测到的大量的蛋白质都是疏水的；可是由二维电泳和质谱技术所检测到的蛋白质却缺少疏水性蛋白质。这表明，二维电泳不能够很有效的展示蛋白质组中的疏水性组分。近年来，Rabilloud 小组通过改进操作方法及化学试剂而在更加有效的提取恢复疏水性蛋白质方面做出了很大的努力（Adessi et al., 1997; Blisnick et al., 1998; Chevallet et al., 1998; Goldberg et al., 1996; Rabilloud, 1998; Rabilloud et al., 1999; Santoni et al., 1999; Tastet, 2003; Luche et al., 2003）。但是，提取疏水性蛋白质仍然是蛋白质组学研究中的一个巨大挑战。关于二维电泳中生物样品的蛋白质提取与溶解方面的综述性文章已有发表（Dunn and Corbett, 1996; Rabilloud, 1996, 2002; Ramagli, 1999）。

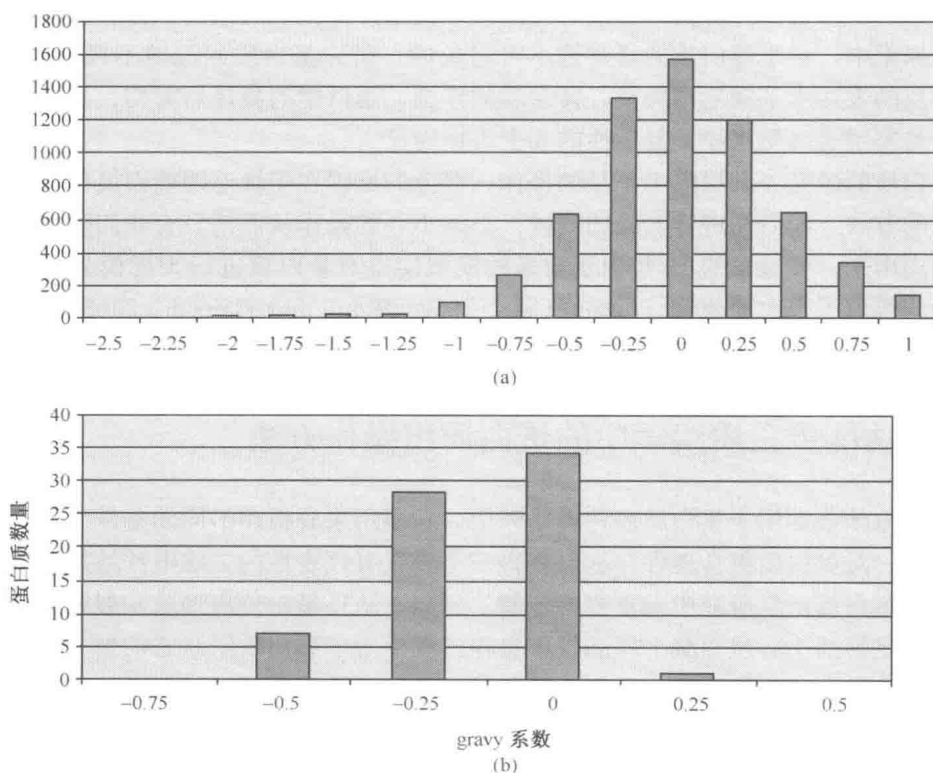


图 1.4 疏水性与可见蛋白质的关系

(a) gravity 系数是根据所有已知的酵母开放阅读框计算获得的。(b) 二维电泳上可见的一些蛋白质的 gravity 系数的计算。数据根据 Gygi et al. (2000) 和 Perrot et al. (1999) 获得。负的 gravity 系数表示为亲水性的蛋白质而正的 gravity 系数表示为疏水性的蛋白质

目前的蛋白质提取方法对于一些特殊的蛋白质组来说通常是行不通的。有的组织包含有许多不同的细胞，而这些细胞又处于不同的状态，每一个细胞都有其特定的蛋白质