

药物蛋白质组学

PHARMACEUTICAL PROTEOMICS

主编 / 郭葆玉



人民卫生出版社
PEOPLE'S MEDICAL PUBLISHING HOUSE

药物蛋白质组学

主 编 郭葆玉

副主编 邱 磊

编者(以姓氏笔画为序)

丁 力	厉建中	郭葆玉
弓雪莲	冯 皓	郭满盈
王 栋	邱 磊	曹颖瑛
王菊蓉	张 冉	道书艳

人民卫生出版社

图书在版编目(CIP)数据

药物蛋白质组学/郭葆玉主编. —北京:人民卫生出版社, 2007. 9

ISBN 978 - 7 - 117 - 09166 - 4

I. 药… II. 郭… III. 蛋白质 - 基因组 - 应用 - 药理学 IV. TQ463 Q51

中国版本图书馆 CIP 数据核字 (2007) 第 132009 号

药物蛋白质组学

主 编: 郭葆玉

出版发行: 人民卫生出版社(中继线 010 - 67616688)

地 址: 北京市丰台区方庄芳群园 3 区 3 号楼

邮 编: 100078

网 址: <http://www.pmph.com>

E - mail: pmph@pmph.com

购书热线: 010 - 67605754 010 - 65264830

印 刷: 尚艺印装有限公司

经 销: 新华书店

开 本: 787 × 1092 1/16 印张: 19.25

字 数: 445 千字

版 次: 2007 年 9 月第 1 版 2007 年 9 月第 1 版第 1 次印刷

标准书号: ISBN 978 - 7 - 117 - 09166 - 4/R · 9167

定 价: 38.00 元

版权所有, 侵权必究, 打击盗版举报电话: 010 - 87613394

(凡属印装质量问题请与本社销售部联系退换)

前 言

基因组研究形成了 20 世纪生命科学研究一道亮丽的风景线，取得了巨大的成就。那么，知道了人类的全部基因组序列，就可以了解和解决各种医学问题吗？其实并不是这么简单。仅凭基因组学这只单脚圆规，很难绘出完美的圆圈。随着人类基因组计划的逐步完成，科学家们又进一步提出了后基因组计划，蛋白质组研究是其中一个很重要的内容。

蛋白质才是在人体内真正发挥作用的物质，它在体内扮演着构筑生命大厦的“基石和砖块”的角色，是生命活动的真正执行者和体现者。基因好比一张制造宏大建筑物的图纸，蛋白质就是根据图纸制造的真正的建筑物。虽然目前已完成了对人类基因的全部序列测定，但是单凭制造建筑物的图纸，不可能演绎出一个建筑群，也不会知道一个建筑群是如何构建、如何进行的。基因的主要功能是通过其表达产物——蛋白质来实现的，而蛋白质亦具有自身特有的活动规律，随着生命活动的进程表现出极其动态的紧密协调的变化，蛋白质在合成之后具有相对独立的修饰、转运和相互间作用能力，同时还具有对外界因素发生反应的能力。因此，只有从蛋白质组学的角度对所有蛋白质的总和进行研究，即开展蛋白质组学研究，才能更加贴近对生命现象和本质的掌握，生命活动的本质和活动规律才能找到答案。正是因为这样，国际科学界预言，在 21 世纪，生命科学的热点将从基因组学转向蛋白质组学，使后者成为新的前沿。蛋白质组学研究不但对了解生命本质有重要的意义，而且在疾病诊断治疗、药物筛选等方面有广泛应用前景，其中蕴藏着开发疾病诊断方法和新药的线索。

现在已经知道，在疾病中只有一小部分是起因于基因突变，而各种疾病都有蛋白质谱的动态变化。每种疾病在不同的发病阶段，在任何症状出现之前，在蛋白质水平方面已经发生了变化。所以通过蛋白质组学研究提供给我们大量的、完整的、动态的蛋白质谱，通过对正常个体及病理个体间的蛋白质组比较分析，我们可以找到某些“疾病特异性的蛋白质分子”，它们可成为新药物设计的分子靶点。通过这些靶点，在药物发现阶段，我们可以对大量新化合物进行自动筛选，缩短新药发现周期；或者通过这些被确认的蛋白质变化标志物，发展成为临床早期诊断指标，为疾病的早期诊断提供分子标志，形成未来诊断学和治疗学的理论和应用基础。蛋白质组的研究不仅是探索生命奥秘的必须工作，也能为多种疾病机理的阐明及攻克提供理论根据和解决途径，为人类健康事业带来巨大的利益。

据统计，在上世纪 90 年代中期，全世界范围内共找寻到的新药药靶约 483 个，其

中大部分是蛋白质；而当时全世界正在使用的共 2,000 多种药物中，绝大部分都是针对上述 483 种药靶的。这 483 种药靶分子构成了全世界制药产业的最重要的发展源头。据研究，按人类主要的 100 多种疾病进行计算，还应该有 3,000 ~ 15,000 种的蛋白质具有开发成为药靶的可能，也就是说还可能有上万种的新药靶将被发现和开发，这些潜在的发展源头，将有可能给制药界带来无穷的财富和发展空间。这也是为什么蛋白质组学作为发现药靶的主要技术平台，越来越受国际制药业界垂青的重要原因所在。

基于上述，我们编写了这本专著，其主要内容涵盖了蛋白质组学理论及基础，蛋白质组学的应用技术、蛋白质组学的数据库及分析方法，同时对蛋白质组学的主要使用工具——二维电泳技术进行了比较详细的介绍。由于在进行药物研究和药学靶点的研究中，蛋白质组学起到了非常重要的作用，所以本书在编写的过程中增加了这方面的内容。本书适合于科研人员和高校的教学人员及研究生的学习参考之用。由于编著的知识水平有限，难免在编写的过程中有很多的缺点和不足之处，敬请读者批评和指正，我们将虚心地接受并在今后的工作中努力克服和提高。

郭蓓玉

2007 年 8 月于上海

目 录

第一部分 基础理论

第一章 概述	3
第一节 基因组研究简介	3
第二节 蛋白质组	5
第三节 蛋白质组学的分类	7
第四节 蛋白质组学研究的用途	11
第二章 功能基因组学与蛋白质组学	18
第一节 人类基因组计划及功能基因组学	20
第二节 蛋白质组研究的开端及含义	35
第三章 双向凝胶电泳	41
第一节 概述	41
第二节 试剂、设备和溶液配制	51
第三节 实验操作	53
第四章 生物质谱	72
第一节 质谱的基本原理	73
第二节 生物质谱的软电离方式	75
第三节 常用质谱质量分析器	79
第四节 质谱法在蛋白质和多肽分析中的应用	82
第五节 质谱法对蛋白质和多肽一级结构的分析和鉴定	86
第五章 蛋白质芯片	93
第一节 概述	93
第二节 蛋白质芯片分析	98
第三节 蛋白质芯片的优缺点及展望	109
第六章 酵母双杂交系统	111
第一节 双杂交和其他双成分系统	111

第二节	酵母双杂交的原理及方法	117
第三节	酵母双杂交系统的发展	120
第四节	酵母双杂交系统的应用	126
第七章	非胶蛋白质组学	136
第一节	一维色谱与质谱联用	136
第二节	多维色谱分离技术	137
第三节	毛细管电泳-质谱联用	140
第八章	亚细胞蛋白质组学	143
第一节	概述	143
第二节	亚细胞结构的分离	146
第九章	生物信息学在蛋白质组研究中的应用	155
第一节	生物信息学简介	155
第二节	常见蛋白质数据库	159
第三节	生物信息学在蛋白质研究中的应用	161
第十章	蛋白质组学在医学中的应用	185
第一节	蛋白质组学在消化系统疾病方面的应用	186
第二节	蛋白质组学在生殖系统方面的应用	189
第三节	泌尿系肿瘤的蛋白质组研究	192
第四节	蛋白质组学在眼科方面的应用	193
第五节	血液系统疾病蛋白质组学研究进展	195
第六节	蛋白质组学在耳鼻咽喉科疾病机理研究中的应用	197
第七节	蛋白质组学在呼吸系统中的应用	199
第八节	医学检验方面的应用	200
第九节	蛋白质组学在其他方面的应用	203
第十节	总结与展望	204
第十一章	蛋白质组学在药学研究中的应用	207
第一节	蛋白质组学在药物作用靶点研究中的应用	207
第二节	药物作用机制研究及新药筛选	218
第三节	蛋白质组学在药物毒理学中的应用	226
第四节	蛋白质组学在中药现代化中的应用	230

第二部分 实际操作步骤

步骤 1	样品的制备	237
------	-------	-----

步骤 2	等电聚焦	247
步骤 3	聚丙烯酰胺凝胶电泳	254
步骤 4	凝胶染色	269
步骤 5	凝胶的扫描和图像分析	273
步骤 6	荧光差异凝胶电泳	278
步骤 7	斑点切除	281
步骤 8	样品褪色	283
步骤 9	凝胶内消化	284
步骤 10	微量脱盐和样品浓缩	284
步骤 11	多肽消化物的化学衍生化	285
步骤 12	质谱分析	286
步骤 13	MALDI-TOF MS 的标定	288
步骤 14	数据库分析的准备	290
步骤 15	PMF 肽指纹图谱数据库检测	291
附录	故障排除	293

第一部分 基础理论

第一章

概述

人类基因组计划 (human genome project, HGP) 是人类生命科学研究中的一个重大里程碑事件, 曾经被人们与曼哈顿计划和阿波罗计划相提并论。在惊奇于生命本质的奥妙的同时, 研究人员也在困惑, 当 HGP 完成以后, 生命科学研究将何去何从。但是随着研究的逐渐深入, 人们逐渐了解到, 通过 HGP 解析出人类基因组的全部序列并非生命科学研究的终结, 而仅仅是一个起点, 在了解了遗传密码以后, 我们似乎有了更多的工作要去做。2003 年 4 月 14 日, HGP 首席科学家 Collins 博士在华盛顿宣布人类基因组序列图谱绘制成功, HGP 的所有目标全部实现, 这标志 HGP 完成, 同时也标志了“后基因组时代”(post genome era, PGE) 的来临。2001 年的 Science 杂志将蛋白质组学列为 2002 年六大研究热点之一, 其“热度”仅次于干细胞研究。

其实, 由于蛋白质在生命活动中的重要性比核苷酸更加直观, 所以人们对蛋白质的研究早于 DNA 和 RNA。最初被破译的蛋白质是肌球蛋白和血红蛋白。出生在维也纳的分子生物学家佩鲁茨, 从 1953 年开始就在卡文迪许实验室用 x-衍射的方法研究蛋白质的结构。4 年后他的学生肯德鲁, 在世界上第一个破译了肌球蛋白的结构。1959 年, 佩鲁茨破译了更为复杂的血红蛋白的结构。1962 年, 他们共同获得了诺贝尔化学奖。对第一个较为复杂的蛋白质结构测定, 科学家整整花了 6 年的时间, 但现在生物化学家给蛋白质数据库增添一个新成员, 仅需几个月乃至几天。科学家至今已经测定并且记载了 8000 多个蛋白质的结构数据。

第一节 基因组研究简介

蛋白质组的概念是在人类基因组计划过程中提出的, 因此在介绍蛋白质组计划之前, 我们先简单介绍一下人类基因组工程。人类基因组工程是由美国科学家, 诺贝尔奖获得者 Renato Dulbecco 于 1986 年在《Science》上发表的论文中最先提出的, 它旨在阐明人类基因组中 3×10^9 个核苷酸的序列, 阐明所有人类基因并确定其在染色体上的位置, 从而破译人类全部遗传信息。1989 年, 布什在白宫的一次国家科学与技术颁奖典礼上, 列举里根政府和布什政府为科学所做出的贡献时提到了太空站、超级超导对撞机、Gnome 等创新。对于总统把单词的拼错, 大家并未表示惊讶, 其中还包括基因拼接技术的发明者 Stanley N Cohen 和 Herbert W Boyer。但是人们认为政府对这项计划不是十

分重视,一般人也困惑于花费如此巨大的经费来进行目前还没有发现有太大用途的研究是否值得。而那一年国立卫生研究所准备为 HGP 的早期研究斥资 2820 万美元,次年正式启动了 HGP。到如今已经不会有人犯这种错误了, HGP 被称为继曼哈顿原子弹计划和阿波罗登月计划后人类科学史上又一伟大工程。

在 HGP 的最初几年里,由于测序技术及测序仪的限制,加上公众的不理解,影响了测序的进度。1993 年,柯林斯开始主管 HGP 后,测序的进度大大提速。1995 年,国家人类基因组计划机构(national human genome institute, NHGI)开始加速 HGP 的研究进展,他们重点资助 6 家从事人类基因组测序的实验室。如今,随着欧共体、日本、加拿大、巴西、印度、中国的加入, HGP 已经成为一个全球范围的合作项目;同时随着数理化、信息、材料等学科的渗透和工业化管理模式的引进, HGP 已经成为生命科学领域的科学工程。同时,由于看到了该研究在商业领域的巨大前景,很多商业机构也积极的加入进来,其中 Celera 公司对 HGP 研究起到了至关重要的推动作用。Celera 公司首席科学家 Venter 提出全新的“鸟枪法”测序的理论,并且用这种思路用一年的时间就完成了 NHGI 几年的工作量,使得 HGP 能够提前完成。

到目前为止, HGP 已经提前完成了人们预期的设想,对人和多种生物的基因组进行了测序,同时将很多成果应用于医药和畜牧业的实践中,取得了一定的成功。在基因组的研究方面,我国根据自身研究经费比较少,而人口分布广泛的基本国情,采取了与国际不完全相同的研究策略。作为一个全球性的合作项目, HGP 采取了全基因组或者整条染色体作图、测序的策略,而中国则突出基因组的多样性和疾病图谱的研究,以局部的识别和测序为主,并且在整个测序的合作中承担了 22 号染色体短臂的序列测定,占整个工作量的 1% 左右。

随着 HGP 的加速运行,人们对 HGP 以后的工作进行了进一步的思考。大多数科学家都认为 2003 年完成的 HGP 只是一个以测序为主的结构基因组学(structural genomics),只是一个开始,那么如何将这些序列信息应用到医药卫生、人类生活质量的提高中去才是关键的问题。于是功能基因组学(functional genomics)就应运而生了。结构基因组学研究的主要对象主要是基因组的序列特征,它是人类对自身认识的一个起步。从中我们可以对基因的复杂度与生物进化、基因组编码序列的结构、基因与蛋白质家族的大小及其进化等问题进行研究。

现在又不断涌现出与基因组相关的新的学科:

人类基因组多样性计划(human genome diversity project, HGDP): 1991 年,一些人类遗传学家和分子生物学家建议对全球各地理区域的人类基因组的多样性进行研究。这一建议得到了人类基因组组织(human genome organization, HGO)的支持,同时也吸引了分子生物学家、人类遗传学家、考古学家、进化学家、语言学家和历史学家的兴趣。1993 年 9 月,举行了人类基因组多样性的工作会议,正式筹备 HGDP 计划,预计 5 年时间,资金 2500 万美元。其目标是利用新技术研究基因的遗传学,结合历史学、考古学和语言学等准确地定义世界上不同人群的起源,其重要性是对人类历史和身份有一个广泛的说明和解释,所建立的资源库为寻找引起或抵抗疾病的遗传因素提供信息。它为自然科学和人类学架起了桥梁,这是其理论研究方面的意义。在实际应用方面,它将服

务于对一些疾病,如遗传流行病的研究。我们知道有些疾病的分布有明显的种族倾向性,如果能了解人类基因组的多样性,就有利于我们找到一些遗传疾病的致病基因,研究出更好的诊断和治疗方法,这是其实际意义。

环境基因组计划 (environmental genome project, EGP): 1998 年 4 月 4 日由美国国立顾问环境卫生委员会正式设立专项基金,预计 6 年间投资 6000 万美元。此计划是在美国人群中选择具有不同年龄、性别和种族背景的 10 个群体中的 1000 个个体,应用 HGP 中所使用的方法,鉴别与环境相关疾病易感基因的等位片段多态性 (allelic polymorphism),建立这些基因多态性的中心数据库。通过对这个数据库的分析,科学家能准确地预测出影响人类健康的环境因素,并帮助政府制定出环境保护政策。现在研究较多的有癌症、呼吸系统疾病、退行性神经系统疾病、发育障碍、先天性缺陷、生殖系统疾病、自身免疫疾病等。

药物基因组学 (pharmacogenomics): 随着生命科学的发展,我们已经知道,基因组的多样性在一定程度上决定人体对药物的反应,通过对影响药物代谢或效应通路有关基因的编码序列进行再测定,揭示个体对药物反应差异的遗传基础,就是“药物基因组学”。药物基因组学不是研究疾病的遗传因素,而是探讨药物作用的遗传分布来满足临床需要,即在做出任何处方决定之前先鉴定患者的相关基因并对患者做试验,再给特定的患者开出恰当的药。

除此以外,人们还将基因组的研究向其他纵深方向拓展开来,产生了很多新兴学科以及学科的交叉。在了解了基因组所传递的遗传信息以后,研究的重点开始转移到这些遗传信息如何在生物个体发育的过程中,如何在机体受到内外刺激的情况下产生相应的反应,复杂生物体内的各个部分又是如何进行分工协调的,而这些工作更加直接的参与者就是蛋白质,于是在基因组研究的基础上又产生了蛋白质组学的概念。

第二节 蛋白质组

虽然基因组的研究向我们展示了生命的众多奥妙,加深了人类对自身和其他物种的了解,并且为改善人类的生活质量做出了重要的贡献,但是它还是给我们留下了一些疑问:为什么酵母的 DNA 测序完成已经有 6 年,到目前为止还无法确定其基因组的准确大小?为什么一些阿米巴虫的 DNA 数目是人类的 200 多倍,而线虫的基因也比苍蝇来的更多?不同人种的基因差异可以达到 0.1%,为什么人与黑猩猩的基因只相差 1%,就会引起如此大的差异?是不是基因的数目与生物的进化之间有直接的联系?如果不是,数量少的基因又是通过何种方式来调控更为复杂的生物个体?这些问题都是基因组研究无法解决的。而问题的关键在于 DNA 虽然在生命活动中起到决定性作用,但是它的作用的发挥是通过一系列的步骤,最终通过蛋白质发挥出来的,而在这个过程中的每一个步骤都可能对最终的结果发生影响 (见表 1-1-1)。

最初提出的一个基因一个蛋白的概念已经得到了修正,事实上在人体中一般一个基因会对应两到三种不同的蛋白质产物,这些变化就是在转录、翻译和翻译后的加工过程中完成的。因此要想了解人体复杂机器如何精密调控,还必须了解细胞中蛋白质产物的

组成及其相互作用，于是针对基因组，人们就提出了蛋白质组的概念。对于一个细胞来说，它所包含的基因组信息是固定不变的，但是蛋白质的组成是有一定的时空性和可调节性的，在不同的细胞内外环境调控下，在细胞或者个体发育的不同时期，细胞内的蛋白质在种类和数量上都有着很大的区别，因此蛋白质组是特定时间、特定环境或实验条件下蛋白质的表达情况，它从整体水平上反映了体内动态的、整体的代谢变化的过程。

表 1-1-1 蛋白质的调节方式

方式	内 容
转录调节	信使 RNA 水平发生变化
质量调节	mRNA 稳定性调节
翻译调节	信使 RNA 水平未变化
蛋白质降解调节	翻译后蛋白质水解处理
	翻译后蛋白质共价修饰
功能调节	调节转运
	蛋白质之间的相互作用

HGP 的发展促进了蛋白质组学的产生，并且为它的研究提供了理论上的基础，但是蛋白质组学研究的高速发展还需要技术和资金上的支持。

在技术方面，蛋白质组研究的两个重要的步骤就是蛋白质的分离以及鉴定。在蛋白质的分离方面，它依赖于二维电泳技术的出现。与传统的一维电泳不同，二维电泳可以利用两种不同的分离方法，将细胞内众多的蛋白质分离，并且使其直观地展示出来，使得对逐个蛋白质或者一些蛋白质复合物的研究成为可能。

在蛋白质的鉴定方面面临的最大问题就是一般生物材料的量都比较有限、二维电泳的上样量也有一定的限制，所以就要有高灵敏度的测序方法。最初解决这个问题是通过 Edman 降解法来测序，后来出现的通过微测序技术对电印迹的样品进行测序是一个重大突破。该技术用于二维电泳的蛋白质分析，从而产生了最早的数据库。九十年代对该技术的改进提高了 Edman 降解法的灵敏度，从而可以在 pmol 水平上对蛋白质测序。

另一个重要的发展就是 MS 技术的发展。过去十年中 MS 用于蛋白质鉴定的灵敏度和准确率都有了很大的提高，现在胶中的 fmol 水平的蛋白质都可以被鉴定。由于 MS 更加敏感、可以处理蛋白质的混合物、可以用于高通量的样品处理，它已经取代了 Edman 降解法的地位。这些技术的发展和成熟为蛋白质组研究的开展提供了可能。

生命科学的研究通常都需要大量资金和人力、物力的投入，因此，资金问题是蛋白质组研究的一个重要难题。生命科学的研究从基因组转移到蛋白质组研究的最大动力就是药物研究需要更快的速度和更准确的方向。以目前的大规模的筛选和化学合成的方法用于药物生产来计算，开发一个新药一般要用 10 ~ 12 年的时间，花费 500 000 ~ 750 000 美元。同时在临床试验阶段由于药物的毒性或者低效等问题也会造成药物无法

在临床使用。现在有望通过蛋白质组的平台来减少药物开发总费用的 30%，缩短 2 年的时间。与基因组在药物研发中的作用相比较，蛋白质组的优势是：①同一个蛋白质组的平台可以用于测试多种生化性质，当在多个测试中的结果都为阳性时，药物靶标选择正确的可能性就大大提高；②利用蛋白质组可以直接检测与疾病相关的生物分子，而基因组的应用则没有这么直接。研究表明，只有约 2% 的疾病与基因序列有关，98% 的疾病与蛋白质的表达有关，仅依靠基因水平的分析有时并不足以了解疾病发生的机制；③用于药物研发的蛋白质组平台还可以直接用于临床前研发和临床后继续研究。因此蛋白质组的研究在药物开发方面有着很大的应用前景。

正是由于意识到了蛋白质在生命活动中的重要地位，以及对蛋白质的研究有可能对医药行业带来的巨大影响和商业方面的无限潜力，美国、英国、日本等国纷纷投巨资资助蛋白质组学的研究。英国科学家试图绘制人类蛋白质图谱，找出与人类疾病有关的蛋白质互动模式。美国能源部（Department of Energy, DOE）在 HGP 之后又启动了“生命基因组计划”（the genomes of life program），主要目的是“识别执行关键生命功能的多蛋白质复合物，分析其基因调控网络特征”。英国生物技术和生物科学研究委员会资助了三个研究中心对一些已经或者即将完成全基因测序的生物进行蛋白质组的研究。在法国也有五个区域性遗传基地得到国家的资助进行蛋白质组方面的研究。德国也建立了专门的蛋白质组研究基地，而澳大利亚在 1997 年就建立了全国性的蛋白质组学研究的网站 APAF（Australia Proteome Analysis Facility）。瑞士的 GeneProt 公司是由以蛋白质组数据库“SWISSPROT”著称的蛋白质组研究人员组成，以应用蛋白质组技术开发新的药物靶标为目的，建立了配备上百台质谱仪的高通量技术平台。曾经在 HGP 中发挥了重要作用的 Celera 公司首席科学家 Venter 博士指出：人类基因组、蛋白质组和药物是生命科学研究的三个阶段，Celera 公司已经进入了蛋白质组研究的阶段。该公司添置大批蛋白质鉴定和分析设备用于蛋白质图谱的绘制。日内瓦蛋白质组公司与布鲁克质谱仪制造公司联合成立了国际上最大的蛋白质组研究中心。此外，还有众多的医药厂家也正准备加入到这个新兴的研究领域中来。正是这些公司建立的巨大数据库网站，使得各国、各个研究机构的研究人员可以及时地将资源和研究成果加以分享。

第三节 蛋白质组学的分类

蛋白质组（proteome）的概念最早是由澳大利亚科学家 Wilkins 和 Williams 于 1994 年提出的，它是由 protein 和 genome 两个词组合而成，1995 年 7 月它首次出现在《Electrophoresis》杂志上。蛋白质组是指一个细胞、组织或有机体所表达的全部蛋白质。随着对蛋白质组研究的深入，它的含义也在不断地扩展。现在，对于蛋白质组学有两种不同的定义：狭义的定义是对研究对象中全部基因表达产物的大规模分析，研究对象仅限于蛋白质；广义的定义是将蛋白质的研究与 mRNA 分析、基因组研究以及酵母双杂交结合。尽管定义的方式不同，它的最终目的都是一样的，就是将细胞中的所有蛋白质作为一个整体加以研究来了解生物学，而不是将各个蛋白质组分单独加以研究。根据更为

广义的定义，很多研究内容都可以归纳到蛋白质学中（见图 1-1-1），例如蛋白质与其他生物活性大分子的相互作用（例如其他蛋白质、脂质、核酸等）、蛋白质的修饰、蛋白质的功能以及蛋白质的定位等等。为了实现对蛋白质在细胞内三维空间上的定位，分子生物学、生物化学、生物信息学等很多领域都将被涉及。对于特定的研究对象来说，蛋白质组学的研究是一个动态的过程，蛋白质的翻译后修饰，细胞内转移、合成或者降解都是由当时细胞内外的环境所决定的。因此，检测细胞内的蛋白质组的状况就可以便捷地了解细胞所处的状态。

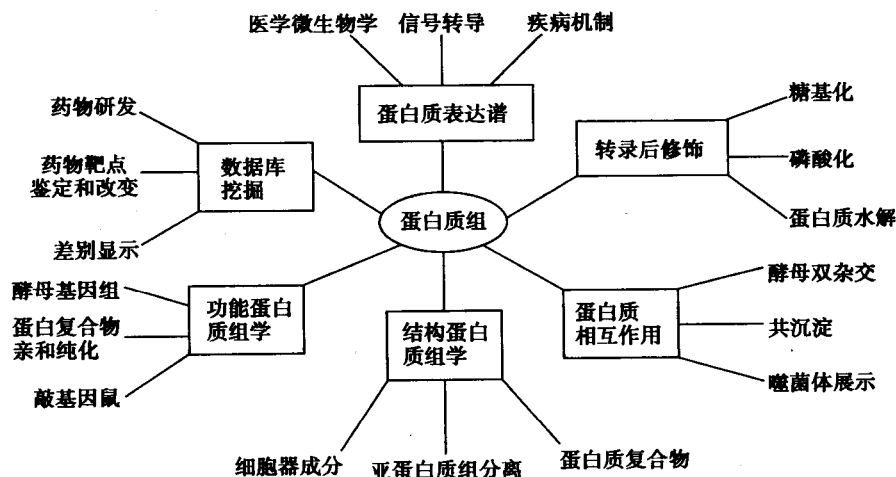


图 1-1-1 蛋白质组学研究的相关领域

目前，蛋白质组学的研究目标如下：建立蛋白质数据库，在今后的十年中，逐步搞清所有的蛋白质分子的结构和功能，并记录在案；建立公用的蛋白质数据网络平台，在分子水平上研究基因对蛋白质合成的调控，特别是蛋白质加工中空间结构的形成过程；蛋白质在生命活动中调控模式的研究，为靶向药物的开发构建实用技术；跨膜蛋白质结构与功能研究及工业、环保上的应用；蛋白质分子结构和功能测定的自动化仪器、软件的开发研究。

与基因组研究不同，蛋白质组的研究有着自身的特点：

1. 对于一个细胞或组织来说，它的基因组组成是不变的，代表了它的全部遗传特征。它在生物的不同组织和部位，在个体发育与细胞生长的不同阶段都是恒定的，而它的蛋白质组成则会随着时间、环境变化、发育阶段等等因素发生变化。这种变化不仅仅表现在细胞表达的种类不同，同种蛋白质的表达量上也存在着很大的变化。因此基因组是静态的，而蛋白质组则是动态的。

2. 随着 HGP 的进行，我们已经知道了各种生物的基因组的大小，下面的工作就是将基因组中的不同核苷酸序列加以鉴定，这种工作相对来说是一个有限的工作。而对于蛋白质组研究来说，由于同一种蛋白质前体经过不同的加工可以成为不同的蛋白质产物，发挥不同的作用，而体内的这种蛋白质的修饰工作又是随着时间、环境等各种因素不停变化的，因此至少在目前阶段看来，蛋白质组的研究工作似乎会成为一个无限的

工作。

3. 对于基因组来说,一个生物的遗传信息与其 DNA 上的核苷酸的排列顺序有关,因此是一个一维的信息。而对于蛋白质来说,一个蛋白质的功能不仅取决于其氨基酸序列,而且还与蛋白质的折叠方式有着密切的关系(蛋白质的空间结构约有 1 万种折叠类型),因此蛋白质提供的是一个三维的信息。同时蛋白质的正确折叠和正确翻译后的修饰还与环境中的其他生物活性物质以及细胞内的各种理化性质密切相关,因此蛋白质的研究应该是环境中的蛋白质的研究。

4. 基因组研究的是一种生物的整体遗传信息,因此它的研究对象是生物的整体,而蛋白质组研究的对象往往范围更小,甚至只是局部细胞器中的一些蛋白质。例如在对肾脏的一些蛋白质加以研究时,尿液或血液中的蛋白质为一组;细胞的内涵体中蛋白为一组;沿肾小管表达的所有 Na^+ 转换体又是一组蛋白质。因此在进行蛋白质组研究时首先就是要确定研究的蛋白质的群。

5. 蛋白质组学与传统的蛋白质研究也有很大的不同。传统生物学对蛋白质的研究只研究蛋白质的分子量、溶解度等性质,而蛋白质组学除了对蛋白质的定性、定量研究以外,更重要的是研究与生命活动相关的一些特征,例如动力学、细胞通路中蛋白质复合物的作用等。

6. 虽然蛋白质组与基因组在自身的特征以及研究方式上面有着一些差别,但是从根本上来说,蛋白质的结构还是由基因来决定的,因此蛋白质组的研究与基因组的研究存在着相辅相成的关系。

正是由于蛋白质组与基因组之间存在着众多的不同之处,同时又存在大量的联系,因此,在进行蛋白质组研究时,可以从以往基因组研究中借鉴到很多东西,同时又有一些手段方法上的不同之处。从研究的手段来讲,我们可以将蛋白质组学的研究分为“蛋白质组技术体系”和“蛋白质组信息学”两大部分。蛋白质组技术体系包括:蛋白质组的样品制备和鉴定、蛋白质相互作用网络的研究技术等;蛋白质组信息学则涉及到蛋白质组数据库的建立,相关软件的应用与开发和生物信息学,即在研究体系建立与不断完善的同时,针对一些重要的生物学问题进行功能蛋白质组研究。首先,要选择有很好研究积累的细胞与生物大分子(蛋白质—基因)的较完整研究系统,利用所建立的相应蛋白质组图谱方法,获取各种重要条件作用前后的同类蛋白质组图谱,并利用计算机图像技术系统分析比较,确定在数量及形式、修饰度或定位等发生变化的蛋白质群体。然后,对重要内外环境因素、基因突变等条件直接作用相关的蛋白质群体进行时相、初级和次级等变化的详细研究分析,从中发现重要的蛋白质群体及其活动规律和关键蛋白质,再深入进行它的组分及其性质的鉴定。由此创建各种细胞的蛋白质组数据库,并进而开展重要的蛋白质结构预测、三维结构和动态结构研究,在分子水平上深入探索其作用模式、功能机制、调节控制及其在蛋白质群体内或与相关生物大分子间的相互作用。

根据研究的方向不同,研究的目的不同,又可以将蛋白质组的研究大致分为以下几个方面:

表达蛋白质组学(expression proteomics):又称为差示蛋白质组学(differential-dis-