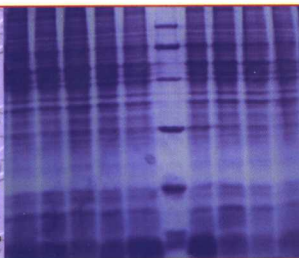
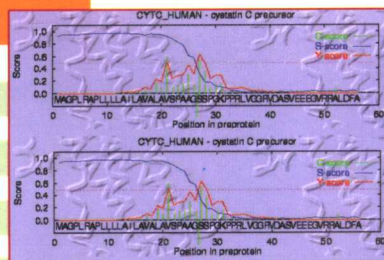


21
世纪生物技术丛书

蛋白质理论与技术

王廷华 邹晓莉 主编



21 世纪生物技术丛书

蛋白质理论与技术

主 编 王廷华 邹晓莉

科 学 出 版 社

北 京

内 容 简 介

蛋白质理论与技术是 21 世纪生物技术发展的一项重要基础实验性科学,是科学研究的重要工具。作为 21 世纪生物技术丛书的一个分册,本书通过上、下篇全面介绍蛋白质的基本理论、进展、技术及其应用。上篇由八章组成,系统介绍了蛋白质的合成、转运、加工与修饰及结构和功能,阐述了蛋白质分离纯化的基本理论,蛋白质定性、定量检测的基本理论,并对蛋白质的生物信息学、蛋白质组学和蛋白芯片理论与进展做了介绍。下篇由六章组成,介绍蛋白质样品的准备,蛋白质电泳技术,层析技术,蛋白质活性及定性、定量检测的方法及应用,并在生物信息学预测蛋白质序列技术方面进行了实验性阐述。

本书可供研究生、本科生和从事蛋白质研究的科研人员阅读,也可作为实验参考用书。

图书在版编目(CIP)数据

蛋白质理论与技术/王廷华等主编. —北京:科学出版社,2005.3

(21 世纪生物技术丛书)

ISBN 7-03-014646-8

I. 蛋… II. 王… III. 蛋白质 - 基本知识 IV. Q51

中国版本图书馆 CIP 数据核字(2004)第 122609 号

责任编辑:吴茵杰 / 责任校对:陈丽珠

责任印制:刘士平 / 封面设计:卢秋红

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 3 月第 一 版 开本: 787 × 1092 1/16

2005 年 3 月第一次印刷 印张:13 1/2 插页:2

印数: 1—3 000 字数: 302 000

定价: 39.80 元

(如有印装质量问题, 我社负责调换<双青>)

《21 世纪生物技术丛书》编审委员会

主 审 蔡文琴 吴良芳 余 哲
委 员 (按姓氏笔画为序)

王廷华	昆明医学院神经科学研究所 四川大学	教授 特聘教授
方秀斌	中国医科大学基础医学院	教授
冯忠堂	昆明医学院神经科学研究所	教授
齐建国	四川大学华西医学中心	教授
羊惠君	四川大学华西医学中心	教授
李云庆	第四军医大学解剖教研室	教授
李建国	复旦大学医学院	教授
李官成	中南大学湘雅医学院	教授
吴良芳	四川大学华西医学中心	教授
吴承远	山东大学齐鲁医院	教授
余 哲	四川大学华西医学中心	教授
应大君	第三军医大学基础医学院	教授
沈馨亚	复旦大学医学院	教授
胡长林	重庆医科大学附属医院	教授
施 静	华中科技大学同济医学院	教授
徐群渊	首都医科大学神经科学研究所	教授
姜保国	北京大学人民医院	教授
曾园山	中山大学中山医学院	教授
蔡文琴	第三军医大学神经科学中心	教授
Jean Philippe Merlio	法国波尔多第二大学医学院	教授
John W. McDonald	美国华盛顿大学	教授
Leong Seng Kee	新加坡国立大学	教授
Eng-Ang Ling	新加坡国立大学	教授
Petrik X Lee	英国伦敦大学	教授
Pierre Dubus	法国波尔多第二大学	教授
Xiong-Zhong Ruan	英国伦敦大学	教授
Xin-Fu Zhou	澳大利亚 Flinders 大学	教授

《蛋白质理论与技术》编委会

主 编 王廷华 邹晓莉

副主编 曾园山 周 雪 吴林艳

编 委 (按姓氏笔画为序)

王廷华 王熙才 王特为 龙双涟

卢永超 冯忠堂 江之泉 李咏梅

杨志敏 杨向东 杨桂枝 邹晓莉

吴林艳 宋新波 张新胜 张 彦

金立德 周 雪 周 健 洪孙权

夏 阳 倪 炜 高志宇 章 为

曾园山

总 序

21 世纪是生命科学取得革命性进展和医学飞速发展的时代。如果说 20 世纪后半叶是信息时代,那么 21 世纪上半叶生命科学将成为主宰。随着我国加入 WTO 和与世界科技日益接轨,生命科学领域里生物技术的竞争已日益呈现出其核心地位和作用。正是在此背景之下,科学出版社组织了这套《21 世纪生物技术丛书》。该套丛书共八本,包括《组织细胞化学理论与技术》、《神经细胞培养理论与技术》、《蛋白质理论与技术》、《分子杂交理论与技术》、《PCR 理论与技术》、《基因克隆理论与技术》、《抗体理论与技术》、《干细胞理论与技术》。

本套丛书从形态、细胞、分子生物学三个层面介绍了 21 世纪常用生物技术的基本理论、进展及其相关技术与应用。从培养科学思维能力和科研工作能力的目标出发,以实用性和可操作性为目的,面向我国日益扩大的研究生招生规模和广大的一线科研人员。在技术章节提供了大量原版彩图及实验经验体会,使丛书更具实用价值。在编写方式和风格方面,力求强调科学史的沿革及基本概念、基本技术和理论的阐述,基本反映了 21 世纪常用生物技术和理论的现状与进展。

丛书由我国青年神经科学专家王廷华教授牵头,邀请国内、外一批知名专家、教授参加编写和审阅。丛书是全体参编人员实践经验的总结,对一线从事科研的研究生和科研人员有较好的参考价值。由于时间有限,加之科学技术发展迅猛,错误、不足之处在所难免,恳请各位前辈、老师、同道及广大读者批评指正。

值本套丛书出版之际,感谢为我国生物技术与科学发展孜孜不倦、奉献一生的老一辈科学家,是他们的指导、培养和杰出工作为我国中青年一代的发展奠定了基础,并为本套丛书提供了参考。感谢国内、外一批知名专家、教授组成的编审委员会对丛书的认真审阅和编者们的辛勤劳动。感谢科学出版社的同志们对丛书出版所付出的辛勤劳动和支持。感谢各位同道给予的鼓励 and 关心。

《21 世纪生物技术丛书》

编审委员会

2004 年 12 月 8 日

目 录

上篇 蛋白质理论

第一章 蛋白质研究历史回顾	(1)
第一节 蛋白质研究的初期阶段	(1)
第二节 蛋白质研究的发展阶段	(2)
第三节 蛋白质研究的分子生物学阶段	(3)
第二章 蛋白质的合成、转运、加工与修饰	(6)
第一节 蛋白质的合成	(6)
第二节 蛋白质合成后的定向输送	(14)
第三节 蛋白质合成后的加工与修饰	(16)
第三章 蛋白质的结构与功能	(19)
第一节 蛋白质结构概念的提出	(19)
第二节 蛋白质的结构生物学	(20)
第三节 蛋白质的基本结构单位	(21)
第四节 蛋白质的一级结构	(25)
第五节 蛋白质的二级结构	(25)
第六节 蛋白质超二级结构和结构域	(29)
第七节 蛋白质的三级结构和四级结构	(31)
第八节 蛋白质的生物学功能	(33)
第四章 蛋白质分离纯化的基本理论	(36)
第一节 蛋白质的理化性质	(36)
第二节 利用溶解度差别分离蛋白质的方法	(38)
第三节 利用分子大小不同的分离纯化方法	(40)
第四节 电泳技术	(41)
第五节 蛋白质化学中的层析技术	(46)
第五章 蛋白质定性定量检测的基本理论	(51)
第一节 免疫组织化学的基本原理	(51)
第二节 蛋白质定量检测原理	(53)
第三节 免疫印迹法分析特定蛋白质的相对含量	(55)
第六章 蛋白质的生物信息学理论	(57)
第一节 蛋白质生物信息学的概念及内容	(57)

第二节	Internet 网上的生物信息学资源	(57)
第三节	序列对比和数据库搜索	(61)
第七章	蛋白质组学理论与进展	(69)
第一节	蛋白质组学的基本概念及历史回顾	(69)
第二节	蛋白质组学研究的技术平台	(71)
第三节	蛋白质组学在医学中的应用与展望	(76)
第八章	蛋白质芯片理论与进展	(80)

下篇 蛋白质技术

第九章	蛋白质样品的准备	(87)
第一节	分离纯化蛋白质样品的方法	(87)
第二节	蛋白质样品的贮存与运输	(91)
第十章	蛋白质电泳技术	(95)
第一节	概述	(95)
第二节	电泳的基本原理	(95)
第三节	聚丙烯酰胺凝胶电泳	(96)
第四节	SDS-聚丙烯酰胺凝胶电泳	(107)
第五节	聚丙烯酰胺凝胶双向电泳——2D 电泳	(110)
第六节	毛细管电泳	(114)
第十一章	蛋白质层析技术	(121)
第一节	层析技术的基本原理及分类	(121)
第二节	柱层析	(121)
第三节	高效液相色谱法	(128)
第十二章	蛋白质的定量检测	(138)
第一节	紫外-可见分光光度法	(138)
第二节	荧光分光光度法	(148)
第三节	HPLC、CE 分离定量	(152)
第四节	酶联免疫吸附试验	(153)
第十三章	生物信息学预测蛋白质序列技术	(160)
第一节	概述	(160)
第二节	蛋白质识别与描述	(161)
第三节	蛋白质序列的物理性质计算	(167)
第四节	蛋白质二级结构和折叠类型分析	(171)
第五节	特殊结构或特征性结构的预测	(178)
第六节	三级结构的预测	(184)
第七节	蛋白质的质谱鉴定	(186)
第十四章	蛋白质技术的应用	(193)
第一节	蛋白质生物活性检测方法之一——用细胞钓蛋白技术检测备用根猫脊髓	

	背角内神经营养因子的生物活性	(193)
第二节	蛋白质生物活性检测方法之二——凝胶小块与背根节联合培养法	(195)
第三节	Western 印迹技术在检测鸡胚脊髓内神经营养因子中的应用	(196)
第四节	牛脑 S-100 蛋白的纯化与鉴定	(200)
彩图		

上篇 蛋白质理论

第一章 蛋白质研究历史回顾

生物学是研究生命现象、生命本质、生命活动及其规律的科学。生物学发展到今天,已经在分子水平、细胞水平和整体水平三个层次上研究生命活动及其规律。生命科学发生巨变,起于20世纪之初。由于化学、物理和数学领域的广泛渗透,给现代生命科学奠定了坚实的基础。特别是1953年, Watson 和 Crick 借助几个实验室的研究成就,根据 DNA 的 X 线衍射图谱,由 A、T、G、C 的物理化学数据,建立了 DNA 双螺旋模型。之后,从20世纪70年代开始,分子生物学逐步形成,使生命科学进入了崭新的阶段,从而在本质上揭示生命活动的真谛。

构成生命活动最重要的物质无疑是蛋白质和核酸,每一生命活动都是由基因表达产物——蛋白质的特定群体来执行。在生命科学探索的长征途中,“蛋白质组”研究的时代已经到来。此外,蛋白质是生物体的基本组成成分之一,也是含量最丰富的高分子物质,约占人体固体成分的45%。生物体内蛋白质的分布较为广泛,几乎所有组织、器官都含有蛋白质。

第一节 蛋白质研究的初期阶段

18世纪中叶至20世纪初是生物化学的初期阶段,主要研究了生物体的化学组成,特别是对氨基酸、糖类及脂类的性质进行了较为系统的研究。化学合成简单的多肽、酵母发酵过程、生物体内的各种化学反应等,几乎都是在特异的生物酶的催化下完成的。

18世纪中期,瑞典化学家 Scheele 研究了生物体各种组织的化学组成。德国著名化学家 Fischer 应用有机化学的方法对生物体内的糖类、脂类、蛋白质和氨基酸等进行了比较详尽的研究,确定了蛋白质是氨基酸通过肽键连接起来的。后来,又成功地用化学方法合成了由18个氨基酸组成的肽。在酶与底物的相互作用上提出了“锁钥”学说,证明了酶催化的高度特异性。德国生理学家 Seyler 成功地从血液中分离出血红蛋白,并制备成结晶。

18世纪中后期,对酶的研究也有了一些进展。酶学知识来源于生产实践,酶的系统研究始于19世纪中叶对发酵本质的探讨。法国著名科学家巴斯德(Pasteur)认为发酵是完整酵母细胞生命活动的结晶。而1897年,德国科学家 Hans Bi chner 和 Eduard Bu chner 兄弟首次用不含细胞的酵母提取液实现了发酵,从而证实发酵过程并不需要完整的细胞,这一贡

献打开了通向现代酶学与现代生物化学的大门。这一阶段主要是客观地描述生物体的化学组成,习惯上称为叙述性生物化学阶段。

值得一提的是,1859年,Charles Darwin 在《On the Origin of Species》一书中创立了物种进化的自然选择学说,这就是著名的达尔文进化论。该学说认为世界上复杂的植物和动物都是由最初的原始生物经过可持续的进化过程衍生出来的,第一次指出了生物性状的可遗传性,在自然选择压力下的可变性以及不同物种之间的相关性。1865年,Mendel 在分析豌豆性状的杂交实验结果时认为,生物体内有某种遗传颗粒或遗传单位,能够从亲代传到子代,这种遗传单位控制着特定的生物性状。1879年,Walter Flemming 在研究细胞分裂时发现染色体。1900年,人们开始把这种控制遗传性状的遗传单位称为基因(gene)。

在我国,古代劳动人民在饮食、营养和医药等方面的创造和发明,也在实践上为生物化学的诞生和萌芽做出了贡献。早在公元前23世纪夏禹时期,仪狄已能做酒,以曲为媒使五谷为酒,就是利用酒母作为媒介物,促进谷物中的糖类转化为酒。公元前12世纪,运用发酵原理制造酱、醋等食品。公元前2世纪汉淮南王刘安会制作豆腐,这说明我们祖先在当时已经会从豆类中提取蛋白质。公元4世纪的晋代,用海藻酒作为医治地方性甲状腺肿的特效药,因为海藻酒含碘比较丰富。公元7世纪,唐代孙思邈则首先用富含维生素A的猪肝治雀目,用富含维生素B₁的车前子、防风、杏仁、大豆、槟榔等治疗维生素B₁缺乏病(脚气病)。公元10世纪起,我国开始用动物的脏器治疗疾病,例如,用紫河车作强壮剂、用蟾蜍治创伤、羚羊角治中风等,可见古人对含内分泌物质的脏器在临床上的应用,已有一定的感性认识。综上所述,我国古代在生物化学的发展上也做出过积极的贡献。

第二节 蛋白质研究的发展阶段

从20世纪初期开始,生物化学进入了蓬勃发展阶段。在物质代谢方面,由于化学分析及同位素示踪技术的发展与应用,对生物体内主要物质的代谢途径已经基本确定,包括糖代谢的酶促反应过程、脂肪酸 β 氧化、尿素合成途径及三羧酸循环等;在内分泌学方面,发现了多种激素,并将其分离、合成;在酶学方面,脲酶结晶获得成功;在营养学方面,发现了人类必需氨基酸、必需脂肪酸及多种维生素。20世纪上半叶,体内各种主要物质的代谢途径均已基本研究清楚,人们将这个时期称为动态生物化学阶段。

1926年,Sumner从刀豆提纯了脲酶,制成结晶纯品,首次证明酶的化学本质是蛋白质。1932年,英国化学家Krebs等提出合成尿素的鸟氨酸循环的多酶反应途径。1937年,Krebs公布了三羧酸循环的研究成果,这是糖、脂类和氨基酸彻底氧化的多酶反应途径。1940年,德国生物化学家Embelen和Megerhof公布了糖酵解途径。随后,脂肪酸氧化和核苷酸代谢等途径也相继被阐明。在这一时期,我国生物化学家吴宪等提出蛋白质变性学说,这是当时最完备的学说。其基本论点,至今为止仍然是正确的。他在血液化学方面创立的无蛋白血液的制备方法及血糖的测定等方法,迄今还为人们所采用。

1902年,Sutton提出了染色体遗传学说,即细胞核内的染色体一般是二倍体。最使人感兴趣的是生殖细胞在减数分裂时,每个配子得到其中一套染色体(单倍体),该学说认为基因是染色体的一部分。1910年,Morgan证实了基因的确存在于染色体上。1944年,Avery

和他的同事们通过实验证实 DNA 是携带遗传信息、构成染色体的生物大分子。

第三节 蛋白质研究的分子生物学阶段

1953 ~ 1970 年,是分子生物学的理论和技术体系逐步形成的时期, Watson 和 Crick 于 1953 年提出的 DNA 双螺旋结构模型,为揭示遗传信息传递规律奠定了基础,是生物化学发展进入分子生物学时期的重要标志。它用分子结构的特征,解释了生命现象最基本问题之一的基因复制规律,从而真正开始了从分子水平研究生命活动。

在这个时期内,先后发现了 mRNA、DNA 聚合酶和 RNA 聚合酶;DNA 半保留复制机制,操纵子学说等先后被提出;遗传密码被发现,其通用性被证明,并于 1966 年破译了全部 64 个密码子。密码子的发现和破译证实了遗传信息的流动方向是 DNA→RNA→蛋白质,这就是生物遗传学的中心法则。中心法则的建立,使分子生物学作为一门科学初步形成了它的理论体系。

DNA 双螺旋模型已经预示出 DNA 复制的规则, Kornberg 于 1956 年在大肠杆菌 (*E. Coli*) 的无细胞提取液中实现了 DNA 的合成。他从 *E. Coli* 中分离出的能使四种 dNTP 连接成 DNA 的 DNA 聚合酶 I, DNA 的复制需要以一个 DNA 作为模板来进行。1958 年, Meselson 和 Stahl 用精彩的实验证明, DNA 复制时 DNA 分子的两股链要先行分开,他们用 ^{15}N 重同位素及密度梯度超速离心证实 DNA 复制是一种半保留复制。

1954 年,物理专家 Gamow 认为遗传信息是由 DNA 分子以含氮碱基非重叠三联体密码来携带的。1961 年, Crick 及其同事用 T_4 噬菌体的实验确认遗传信息是以非重叠的碱基三联体线性顺序的形式而存在的,三联体之间不存在间隔物。Nirenberg 和 Matthaei 以及 Knorand 等以后的工作最终破译了遗传密码。

早在 1953 年, Zamecnik 及其同事在无细胞系统中发现蛋白质的合成场所为核糖体 (ribosome), 他们还证明蛋白质的合成还需要 ATP 作为肽链形成的能源。1960 年以后, 他们利用 T_4 噬菌体感染 *E. Coli* 作为系统, 发现一种 RNA 能携带 DNA 信息并将其转移到核糖体上合成蛋白质, 故称其为信使 RNA (mRNA)。Hurwitz 等发现 RNA 聚合酶, 这种酶以 DNA 为模板利用 ATP、GTP、CTP 和 UTP 等合成 RNA, 这个过程称为 DNA 转录。

在遗传信息决定蛋白质结构方面, 是 DNA 的碱基序列决定 mRNA 核苷酸信息顺序 (转录), 而 mRNA 的核苷酸顺序又决定多肽的氨基酸序列 (翻译), 这一学说称之为中心法则。同时, Temin 和 Baltimore 还发现反转录酶, 证实了从 RNA 到 DNA 的相对于中心法则的逆向信息流的存在, 从而补充了中心法则的完整性。

不同的体细胞有相同的基因组, 但它们所合成的蛋白质种类并不一定相同。从受精卵增殖而来的高等动物细胞, 它们的发育、生长和分化过程受到遗传信息的调控。1961 年, Jacob 和 Monod 提出的 *E. Coli* 乳糖操纵子学说为人们理解动物细胞基因表达调控提供了理想的模型。Sanger 发现胰岛素中氨基酸的排列序列, 显示多肽链中氨基酸残基的特定序列与蛋白质分子的空间结构的关系。1965 年, 我国王应睐等用化学方法合成有生物活性的牛胰岛素, 实现了世界上首次人工合成蛋白质的梦想。

总之, DNA 双螺旋模型的建立, 遗传密码的破译和乳糖操纵子学说的提出是 20 世纪生

生物学界的巨大成果。

1970 年以后,分子生物学飞速发展,理论体系和技术体系不断扩展。近 20 年来,几乎每年的诺贝尔医学和生物学奖以及一些诺贝尔化学奖都授予从事生物化学和分子生物学的科学家,这个事实本身就足以说明生物化学和分子生物学在生命科学中的重要地位和作用。

20 世纪 70 年代初,随着限制性内切酶的发现和 DNA 分子杂交技术的建立,分子生物学进入技术化时代,基因工程学也有所发展。1970 年, Khorana 首次在试管内合成了基因。1972 年, Berg 首次将不同的 DNA 片段连接起来,并且将这个重组的 DNA 分子有效地插入到细菌 DNA 之中。重组的 DNA 进行扩增,于是产生了重组 DNA 克隆。1973 年,重组 DNA 技术的问世极大地推动了人类遗传学和医学遗传学的发展。1977 年, Sanger 应用酶学技术法和 Gilbert 采用化学方法建立 DNA 测序方法,使我们得以了解到基因甚至基因组的结构。到 1977 年,美国成功应用基因工程方法生产出生长激素抑制素,这一突破性的成果震撼了全世界。

1977 年, Shine 克隆人类第一个基因。Kan (1976 年)、Wong (1978 年) 和 Dozy 等 (1979 年) 应用 DNA 实验技术对胎儿羊水细胞 DNA 做检测, 得出了 α -地中海贫血出生前的诊断。1978 年, Kan 等利用胎儿羊水细胞也得出了镰状红细胞贫血出生前诊断, 从而使得许多遗传性疾病能在 DNA 水平上得出诊断。1986 年, Mullis 等建立的聚合酶链式反应 (polymerase chain reaction, PCR) 技术, 使分子生物学的发展进入了一个崭新的阶段。1986 年, Cech 发现了核酶 (ribozyme), 表明 RNA 除具有原先人们认识的功能外, 还具有催化功能。20 世纪 90 年代初, 基因治疗 (gene therapy) 进入临床试验阶段。值得一提的是, 1983 年王德宝等用化学方法合成了有生物学功能的酵母丙氨酸转运核糖核酸, 标志着我国生物化学的研究在某些方面已经达到世界领先水平。

1990 年, 开始实施的人类基因组计划 (human genome project) 是生命科学领域有史以来最庞大的全球性研究计划, 它将确定人基因组的全部序列, 以及人类约 10 万个基因的一级结构。人类基因组 (human genome) 指的是单倍体染色体 (22 条染色体以及 X 和 Y 性染色体) 中所含的全部连锁群 (linkage groups) 的基因, 它包含一套完整的人类基因。人类基因组计划对生物进化的研究、疾病机制的阐明和生物学基础研究均具有重大的意义, 并且对疾病基因的克隆具有很大的潜在商业价值, 因此, 引起了全世界的极大关注。美国、欧共体、意大利、日本、俄罗斯、法国、加拿大、澳大利亚和中国相继启动人类基因组研究。

人类基因组研究是人类史上伟大的科学工程, 它的完成是解决 30 亿对核苷酸的物理位置, 但这仅是研究的开始。而弄清楚人类基因组的功能与它所表达的时空调控机制以及解释不同序列的意义, 则将使研究进入高级阶段, 即人类的基因组学研究。

在生命科学探索的长河中, “后基因组学” 的时代已经到来, 它将是 21 世纪生物学上最重要的内容。1997 年, 英国 “多莉” 羊的诞生足以说明 “后基因组学” 时代的提前到来。正如 19 世纪末期, 近代自然科学始于物理学的革命, 21 世纪自然科学的大转变, 将始于生物学的革命。

(高志宇 王廷华 杨志敏)

参 考 文 献

- 冯作化. 2001. 医学分子生物学. 北京: 人民卫生出版社
- 来茂德. 1999. 医学分子生物学. 北京: 人民卫生出版社
- 李亚娟, 李 荣, 阎宏山. 1999. 生物化学. 北京: 人民军医出版社
- 周爱儒. 2000. 生物化学. 北京: 人民卫生出版社

第二章 蛋白质的合成、转运、加工与修饰

第一节 蛋白质的合成

生物细胞以基因的 DNA 链为模板,以 NTP 为原料,在 RNA 聚合酶的作用下,按照 A ~ U 和 G ~ C 的碱基配对原则生成 mRNA、tRNA、rRNA 的过程,称为 DNA 转录。其中,mRNA 是指导蛋白质合成的直接模板。基因表达的最终产物是蛋白质、tRNA 和 rRNA。

一、信 使 RNA

多肽合成的模板是信使 RNA (messenger RNA, mRNA)。mRNA 的发现是分子生物学发展中的一件重大事件。由于 mRNA 在细胞总 RNA 中所占的比例很小,因此,在过去还没有建立合适的分离技术时,很难把它分离出来。mRNA 的概念首先是从理论上提出来的,然后再用实验给予证实。Jaiob 和 Monod 早在 1961 年就提出 mRNA 的假设。他们认为,蛋白质是在胞质中合成的,但是编码蛋白质的信息载体 DNA 却在胞核内,因此,必定有一种中间物质用来传递 DNA 上的信息。他们对这种中间物质的性质做了如下的预言:

- (1) 信使是一种多核苷酸。
- (2) 信使的碱基组成应与相应的 DNA 的碱基组成相一致。
- (3) 信使的长度应是不同的,因为由它们所编码的多肽链的长度是不同的。
- (4) 在多肽合成时,信使应与核糖体做短暂的结合。
- (5) 信使的半衰期很短,所以信使的合成速度应该是很快的。

mRNA 的假设提出后,还必须用实验来证明这种假设是否正确。Brenner 等设计一组实验,用噬菌体 T₂ 感染大肠杆菌后,发现几乎所有在细胞内合成的蛋白质都不再是细胞本身的蛋白质,而是噬菌体所编码的蛋白质。噬菌体感染后不久,大肠杆菌内出现了少量半衰期很短的 RNA,它们的碱基组成与噬菌体 DNA 是一致的。

他们通过研究发现,噬菌体感染大肠杆菌后并没有引起大肠杆菌内出现新的核糖体,但出现另一些新类型的 RNA,其代谢速度极快。后来,Spiegelman 又用分子杂交技术证明经噬菌体感染后的新合成的 RNA 可以与噬菌体 DNA 相杂交,但细胞内的其他类型 RNA 则不能与噬菌体 DNA 杂交。从而证明新合成的 RNA 是由噬菌体 DNA 编码的,其序列与之互补的 mRNA。

二、遗传密码

据认为天然蛋白质有 $10^{10} \sim 10^{11}$ 种,但是组成蛋白质的氨基酸却只有 20 种。这 20 种氨

基酸排列组合的不同,则形成各种各样的蛋白质。然而蛋白质中氨基酸排列顺序又是如何决定的呢?

如果 mRNA 采用每三个相邻碱基为一个氨基酸编码,则四种碱基对的组合($4^3 = 64$),可以满足 20 种编码的需要。所以这种编码方式的可能性最大。应用生物化学和遗传学研究技术,已经证明就是三个碱基编码一个氨基酸,所以称它为三联体密码(triplet code)或密码子(codon)。1965 年, Nirenberg 等经过 4 年的研究,将遗传密码完整地编排在下列表中(表 2-1)。

表 2-1 遗传密码与氨基酸的关系

第一个核苷酸(5')	第二个核苷酸				第三个核苷酸(3')
	U	C	A	G	
U	苯丙氨酸	丝氨酸	酪氨酸	半胱氨酸	U
	苯丙氨酸	丝氨酸	酪氨酸	半胱氨酸	C
	亮氨酸	丝氨酸	终止密码(赭石型)	终止密码(蛋白石型)	A
	亮氨酸	丝氨酸	终止密码(琥珀型)	色氨酸	G
C	亮氨酸	脯氨酸	组氨酸	精氨酸	U
	亮氨酸	脯氨酸	组氨酸	精氨酸	C
	亮氨酸	脯氨酸	谷氨酰胺	精氨酸	A
	亮氨酸	脯氨酸	谷氨酰胺	精氨酸	G
A	异亮氨酸	苏氨酸	天冬酰胺	丝氨酸	U
	异亮氨酸	苏氨酸	天冬酰胺	丝氨酸	C
	异亮氨酸	苏氨酸	赖氨酸	精氨酸	A
	蛋氨酸	苏氨酸	赖氨酸	精氨酸	G
G	缬氨酸	丙氨酸	天冬氨酸	甘氨酸	U
	缬氨酸	丙氨酸	天冬氨酸	甘氨酸	C
	缬氨酸	丙氨酸	谷氨酸	甘氨酸	A
	缬氨酸	丙氨酸	谷氨酸	甘氨酸	G

在表 2-1 的 64 个密码中,61 个密码分别代表各种氨基酸,一种氨基酸少的只有 1 个密码,多的可有 6 个,但以 2 个和 4 个的居多。UAA、UAG 和 UGA 为肽链的终止密码子(terminator codon),不代表任何氨基酸。琥珀(amber)、赭石(ochre)和蛋白石(opal)分别为三个终止密码的特殊名称。UAG 作为终止密码子,是由 Bernstein 等发现的,琥珀是德语 Bernstein 的词义,因而 UAG 又称为琥珀型无意义密码子(即终止密码子)。密码 AUG 不仅代表蛋氨酸,还因为它位于 mRNA 起动部位,是肽链合成的起动密码子(initiator codon),因此, AUG 称起始密码。遗传密码的特点如下:

(一) 遗传密码的连续性(commaleless)

密码的三联体是不间断的,须三个一组连续读下去。mRNA 链上碱基的插入或缺失,可造成框移(frame shift),使下游翻译出的氨基酸完全改变。

(二) 通用性(universal)

从最简单的生物(如病毒),一直到人类,在蛋白质的生物合成中都使用同一套遗传密码,这一点为地球上的生物是来自同一起源的进化学说提供了有力的依据。但近年来发现在哺乳动物线粒体的蛋白质合成体系中有例外,如 UAG 不代表终止密码字而是代表色氨酸,CUA 不代表亮氨酸而是代表苏氨酸,AUA 不代表亮氨酸而是代表蛋氨酸。

(三) 简并性(degeneracy)

遗传密码中,除色氨酸和蛋氨酸(甲硫氨酸)仅有 1 个密码子外,其余氨基酸有 2 个、3 个和 4 个或多至 6 个三联体为其密码。有两个以上密码的氨基酸,三联体上一、二位碱基大多是相同的,只有第三位不同。这些密码子第三位碱基如出现了突变,并不影响其翻译的氨基酸。遗传密码的简并性是指密码子上第三位碱基的变化往往不会影响原有氨基酸的翻译。同一氨基酸有多个密码子,其中会有一两个是被优先选用的,称为翻译过程对密码子的“偏爱性”。

(四) 摆动性(wobble)

翻译过程氨基酸的正确加入,需靠 mRNA 上的密码与 tRNA 上的反密码子相互以碱基配对辨认。密码子与反密码子配对,有时会出现不遵从碱基互补配对的规律,称为遗传密码的摆动现象,这一现象更常见于密码子的第三位碱基对与反密码子的第一位碱基之间,两者虽不严格互补,也能相互辨认(表 2-2)。

表 2-2 密码子、反密码子配对的摆动现象

tRNA 反密码子碱基	I	U	C
mRNA 密码子碱基	A, C, U	A, G	C, G, U

(五) 遗传密码的不重叠性(non-overlapping)

假设 mRNA 上的核苷酸序列为 ABCDEFGHIJKL……按不重叠规则读码时应读为 ABC, DEF, GHI, JKL, ……每三个碱基编码一个氨基酸,碱基的使用不发生重复。如果按完全重叠规则读码时,则应该是 ABC 编码氨基酸 1, BCD 编码氨基酸 2, CDE 编码氨基酸 3……。

目前已经证明,在绝大多数生物细胞中基因的读码规则是不重叠的。但是在少数大肠杆菌噬菌体(如 R₁₇、QB 等)的 RNA 基因组中,部分基因的遗传密码却是重叠的。

三、核糖体——蛋白质合成的场所

核糖体(ribosome)是一种非膜性细胞器,由核糖核酸和蛋白质组成,故又称为核糖核酸蛋白体。核糖体普遍存在于原核细胞和真核细胞内,是细胞合成蛋白质的重要场所。在具有分泌功能的细胞中,核糖体的数量随着细胞的功能状态而变化,因此,又有动质(ergastoplasm)之称。1955 年,Palade 在动物细胞中首先看到了这种颗粒,并称之为 Palade 颗粒。