

遗传密码

〔英〕B. F. C. 克拉克 著

钟安环 译

科学出版社

1984

内 容 简 介

本书从介绍蛋白质、核酸的结构开始,叙述了如何用生物化学、分子遗传学和化学方法阐明遗传密码的情况,以及大分子组分怎样在细胞内聚集成蛋白质的一些最新知识。

本书可供中学生物老师、大专院校生物系师生、农学院、医学院有关人员参考。

Brian F.C. Clark
THE GENETIC CODE
Edward Arnold 1977

遗 传 密 码

〔英〕B. F. C. 克拉克 著

钟安环 译

责任编辑 高小琪

科 学 出 版 社 出 版

北京朝阳门内大街137号

中国科学院印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1982年11月第 一 版 开本: 787×1092 1/32

1984年6月第二次印刷 印张: 3

印数: 10,151—15,550 字数: 83,000

统一书号: 13031·2053

本社书号: 2806·13—10

定价: 0.50元

总 序

当前要使一本教科书既能概括整个生物学领域，又能充分反映其最新成果，这已经是不再可能的了。同时，中学和大专院校的师生们还需要掌握这个学科的最新动向和了解哪些领域有了重大发展。

为了满足进一步探求这些知识的需要，几年来我们生物研究所主持编辑了这套小丛书，题目由专门编辑小组精心选定，并受到中学和大专院校师生们对这套小丛书的热情欢迎，这就表明这套书的选题范围，特别是在研究方面和观点的进展方面，以及阐述简明而内容新颖，对读者是具有实用价值的。

这套小丛书的特点是注意研究方法，并尽可能为实际工作提出建议。

作者和本研究所主管教育负责人欢迎读者批评。

生物研究所 伦敦

前 言

对遗传密码的阐明，可能是本世纪生物学中的最重大的进展。一切生物的遗传物质内都含有遗传密码的信息，以保证生物遗传的连续性，并且控制生命所必需的代谢过程。早在本世纪五十年代，人们已认识到蛋白质不是遗传物质，而是基因的产物，遗传物质是 DNA，因而开始研究这两种相关的聚合物所含的遗传信息之间的联系。这种关系就是遗传密码。

从1960年到1966年，人们应用分子生物学方面的生物化学、分子遗传学和化学方法，基本上弄清了生物学上这个中心问题的主要梗概。当人们认识到解决遗传密码的目标有可能达到时，他们竞相努力，使这个新学科的研究活动达到了顶峰。人们在这个领域里的研究工作已经取得很多重要成果，从而丰富了我们关于蛋白质生物合成的分子机制的知识。本书包括大分子组分如何在细胞内聚集成为蛋白质的一些最新知识。由于生物学是一门实验科学，所以我试图向读者介绍一些实验细节，以阐明这方面的知识。期望这种方法能帮助读者更好地了解我们所研究的分子生物学的本质和范围，从而对这个题目有更实际的了解。

克拉克

1977 丹麦

目 录

总序

前言

1	细胞的大分子	1
1.1	开场白	1
1.2	分子组织化的水平	1
1.3	细胞组分	6
1.4	蛋白质	7
1.5	核酸	14
1.6	遗传物质	19
2	遗传密码的阐明	25
2.1	中心法则	25
2.2	共线性	26
2.3	遗传密码的性质	27
2.4	遗传学的实验	30
2.5	生物化学的实验：无细胞系统	33
2.6	无细胞系统的离析	37
2.7	人工合成多核苷酸的应用	39
2.8	三联体结合的测定	43
2.9	利用已知顺序的人工合成多核苷酸	46
2.10	全部密码的完成：它的特性和普遍性	49
2.11	摆动假说	51
2.12	解读遗传信息的方向性	52
2.13	遗传密码转译的准确性	54
3	蛋白质生物合成的组分	55
3.1	引言	55

3.2	氨基酸激活酶	55
3.3	适配器(tRNA)	60
3.4	核糖体	67
3.5	mRNA	73
3.6	蛋白质生物合成的循环图式	76
4	标点	79
4.1	蛋白质合成的起始步骤	79
4.2	起始信号	82
4.3	终止作用信号	83
4.4	蛋白质合成的终止步骤	85
4.5	结论	86

细胞的大分子

1.1 开 场 白

验明 DNA 是遗传物质并且确定它的双螺旋结构，可以认为是分子生物学这门新学科发展的开端。沃森 (Watson) 和克里克 (Crick) 提出的 DNA 双螺旋结构，很容易说明如何复制 DNA 以保证在细胞分裂过程中遗传的连续性。然而，十年之后才确定了 DNA 是怎样通过给细胞酶合成提供信息来调控细胞的代谢过程。当本世纪五十年代初期，人们认识到蛋白质是基因的产物而不是遗传物质时，就开始研究这两种聚合物(蛋白质和核酸)语言之间的联系。这种联系——遗传密码，在某种意义上是通向分子生物学的关键。遗传密码的阐明是二十世纪主要的科学成就之一，并且在几个研究中心展开了热烈的争论(表 1)。在详细地说明密码的解释和密码表达时所使用信号的最新知识以前，我先谈谈蛋白质和核酸的结构以及它们在生物系统中的地位，这将是有益处的。

1.2 分子组织化的水平

细胞是生物体的基本单位。细胞能够象大肠杆菌那样自

表1 阐明遗传密码的重大事件

发 现	时 间	研 究 者	工 作 单 位
DNA转化细菌	1944	艾弗里(O. T. Avery), 麦克劳德(C. M. MacLeod)和麦卡蒂(M. McCarty)	洛克菲勒研究所 纽约
DNA是病毒的遗传物质	1952	赫尔希(A. D. Hershey)和蔡斯(M. Chase)	冷泉港实验室 美国
DNA双螺旋模型	1953	沃森(J. D. Watson)和克里克(F. H. C. Crick)	剑桥大学卡文迪什实验室
蛋白质的一级结构	1954-55	威尔金斯(M. Wilkins)	皇家学院 伦敦
适配器假说	1957	桑格(F. Sanger)	剑桥大学生物化学系
tRNA是适配器的发现	1958	霍格兰(M. B. Hoagland)	剑桥大学卡文迪什实验室
作为蛋白质合成模板的核糖体	1958-59	沃森和特赛尔斯(A. Tissieres)	麻省普通医院 波士顿
信使RNA的假定	1960-61	雅各布(F. Jacob)和莫诺(J. Monod)	哈佛大学 美国
无细胞系统的转译作用	1961	格罗斯(F. Gros)	巴斯德研究所 巴黎
细菌和噬菌体遗传密码的性质	1961-62	布伦纳(S. Brenner)和梅塞尔松(M. Meselson)	哈佛大学 美国
蛋白质合成的方向	1961-62	尼伦伯格(M. W. Nirenberg)和马太(H. Matthaei)	加利福尼亚理工学校 美国
多核糖体	1963	奥乔阿(S. Ochoa)和同事们	国家卫生研究所 马里兰州
基因和蛋白质的共线性	1964	克里克, 布伦纳和同事们	纽约大学医学院
		定兹斯(H. Dintzis)	MRC*分子生物学实验室 剑桥
		里奇(A. Rich)和同事们	MIT**坎布里奇 美国
		吉雷尔(A. Gierer)	MIT 坎布里奇 美国
		亚诺夫斯基(C. Yanofsky)和同事们	普朗克研究所 吐本根 德国
		布伦纳和同事们	斯坦福大学 加利福尼亚 美国
			MRC分子生物学实验室 剑桥

三联体结合的测定法	1964	尼伦伯格和李德(P. Leder)	国家卫生研究所 马里兰 美国
一种特异的起始tRNA	1964	马克卡(K. A. Marcker)和桑格	MRC分子生物学实验室 剑桥
一定顺序的mRNA的转译	1965	柯拉纳(H. G. Khorana)和同事们	威斯康辛大学 美国
tRNA的一级结构	1965	霍黎(R. W. Holley)和同事们	康奈尔大学 美国
起始密码子	1965—66	蔡乔(H. G. Zachau)和同事们	科隆大学遗传研究所 德国
终止密码子	1965—66	克拉克(B. F. C. Clark)和马克卡 布伦纳和同事们	MRC分子生物学实验室 剑桥 MRC分子生物学实验室 剑桥
摆动假说	1966	加伦(A. Garen)和同事们 克里克	耶鲁大学 美国 MRC分子生物学实验室 剑桥

* MRC是分子研究中心的缩写。

** MIT是麻省理工学院的缩写。——译者注

由地生活，亦可与邻近的细胞组成特殊的组织(如在哺乳动物组织和器官中那样)。人们不知道是什么东西给细胞以生命，但是现有的全部科学的证据都说明了细胞是由无生命的分子组成的。它的将近30%的质量是由具有一系列亚基结构的大分子组成。这些大分子在细胞中的功能是作为结构部件、能量贮存、遗传信息库和对维持细胞生活的化学过程进行控制的特殊分子。化学家和生物化学家对于这些大分子的物理学和化学的了解远不如对它们组分的了解，但是由于我们克服了处理大分子的困难，因而对于这些大分子的了解就大大增加了。诚然，我们现在已经相当了解蛋白质和核酸的结构，使得我们能够以明确的化学术语来谈论它们在阐明遗传密码时的相互关系。在认识的现阶段，关于结构概念我们可以提出两点概括：第一，大分子的物理性质和化学性质并不仅仅是亚基性质的总和，由于邻近基团和空间结构的相互作用，这些性质是能够发生变化的；第二，就我们所知，生物的大分子和分子都服从物理学和化学的规律，因此，我们无需用新的观念去研究它们的结构和功能。

图1-1列出了一般公认的组成一个细胞分子组织的各级水平。最初从环境中获得一类简单的化合物(如二氧化碳和水)，进而由小分子量(50—250)¹⁾的中间物转变成大分子的基本组件。分子量为100—350的各种亚基分别聚合成它们的大分子(见图)。简单的基本组件(或单体)以共价键分别连接成它们的多聚大分子。这些大分子的分子量变动范围很大(从 10^3 — 10^6)。脂类的分子量比其他大分子低得多(只到2500)，但是它们常常被看作是大分子，因为脂类在大分子装配(例

1) 分子量往往以道尔顿(daltons)这一术语来表示，可简写为d，或全部略去，如上文所示。一个道尔顿等于一个氢原子的质量，即 1.67×10^{-24} 克。

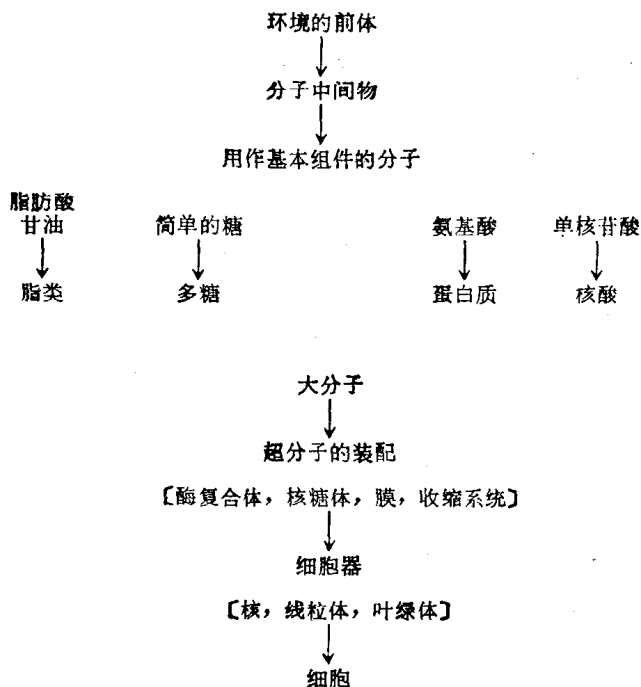


图1-1 分子组织化的水平

如细胞膜)中担负着主要的角色。一方面蛋白质和核酸之间有重要的区别, 另一方面碳水化合物和脂类之间也是如此。根据核酸和蛋白质的结构特点, 人们认为它们是含有信息的大分子。大多数核酸含有四种不同类型的亚基, 而蛋白质则有20种不同的亚基。这些亚基排成线性的序列, 含有遗传信息。然而, 核酸才是遗传物质或遗传的中介物, 而蛋白质则只是基因的产物。的确, 现在基因的正式定义是核酸中的一段, 这一段带有通常为蛋白质以及其他核酸编码的一个遗传信息单位。与这种信息大分子相反, 无信息的多糖和脂类只含有重叠的相同亚基。

在各种细胞中，这四种主要类型的大分子都有同样的功能。核酸贮存和传递遗传信息，蛋白质是基因作用的直接产物或效应物，并含有遗传信息。大多数蛋白质具有特殊的催化活性(例如功能酶)，而另一些蛋白质则作为结构的成分。蛋白质是最富有柔韧性的大分子，所以它具有许多其他的生物学功能，而多糖和脂类在细胞中只作为贮存燃料的成分和结构的成分。

大分子显然不需要新的遗传信息或外加的影响，就能自身组成超分子的装配。象核糖体(核蛋白)那样的超分子装配是如此之大，以致它们在细胞中成为不可溶的颗粒。蛋白质和脂类的复合体也参与有组织的结构(例如膜)；它们的分子量可从 10^8 — 10^9 。超分子装配的一个重要的特点，就是它们不以共价键连接而是以无数弱键(例如弱的静电氢键或非极性(疏水)相互作用)，以及与某种立体化学的契合类似于“手在手套中”这样的空间关系维持着的。这表明大分子能很容易解离成为它的组分。这些非共价键的相互作用的类型，大概是形成细胞中大分子结构和相互作用的基础。

1.3 细胞组分

一般说来，在组成细胞的分子中，它们的各种类型之间是没有什么根本变化的。表2所列的细菌细胞的成分，反映出大多数细胞的通常组分。当然也有例外，特别是那些具有非常特殊功能的细胞和形成有机体的非生命部分的细胞，例如脂肪贮存物或骨骼，头发和羽毛。除了水(它使大多数的细胞组分都处于溶液中)之外，最丰富的细胞成分是大分子，其中蛋白质是细胞组分的最大的一类。蛋白质的功能是多种多样的，它的分子量变化也很大，在表2中所列出的数字从

表2 一个典型的细菌细胞的组分

分子组分	组分的亚分	总含量	平均分子量	不同种类的数目
蛋白质		15	$4-10 \times 10^4$	3000
核 酸	DNA	1	2.5×10^6	1
	RNA	6		
	mRNA		10^6	1000
	tRNA		2.5×10^4	~50
	23S rRNA		1.1×10^6	1
	16S rRNA		5.5×10^5	1
	5S rRNA		4×10^4	1
碳水化合物		3		~50
脂 类		2		~40
小有机分子		2		~500
无机离子		1		10—20
水		70	18	1

40,000到100,000,是大多数蛋白质的平均分子量。按照核酸的结构和功能可以再亚分(如表2所列的缩写式)。这在第1.5节中将有更详细的叙述。其他细胞组分,如碳水化合物和脂类,则与遗传信息的传递毫无关系。

1.4 蛋白质

蛋白质是活的有机体的主要成分。的确,蛋白质的结构和功能是这样的复杂,以致可以把它看作是生命本身的一个属性。蛋白质的功能主要有五种:它可以作为酶、抗体、结构成分、运输工具和代谢的调节者。作为酶的蛋白质具有如下催化性质:控制着有机分子(代谢作用)的合成和降解,并且从进来的分子中释放出能量为基本的生命过程提供动力。抗体是巨大而且复杂的蛋白质亚基的装配体,它保护有机体,防御外物入侵。作为结构成分的蛋白质,规定和维持细胞的构

造。同时要指出，在简单的细菌病毒(噬菌体)中，遗传物质的外壳结构是由蛋白质组成的。病毒被看作是生命的界线。噬菌体除非感染细菌细胞，否则不能称之为生命，因为噬菌体只有在细菌内才能繁殖。作为运输工具的蛋白质，在细胞内或者透过细胞膜传送小分子或离子。最后，作为代谢调节者的蛋白质，能协调和指导细胞内的化学过程，它既可以作为激素，间接作用于遗传信息，也可以作为阻遏物分子，直接作用于遗传信息。很明显，蛋白质与细胞内所有的生物过程都有密切的关系。由于蛋白质的功能是如此多样，所以可以想象，蛋白质化学是十分复杂的，但是目前有关蛋白质结构的知识已足以解释这种多样性。蛋白质不含有传递给其他系统形式的遗传信息；蛋白质只是遗传信息的最终产物。蛋白质的很多功能是短暂的，而且蛋白质的分子本身在细胞中能被迅速地合成或降解，由于可能会出现太多的失误，所以这不合乎作为遗传物质的携带者所要求的性质。于是，遗传物质的角色就由十分稳定的DNA分子来承担了。关于DNA分子的结构将在第1.5节中论述。贮存在DNA分子中的遗传信息必须能在蛋白质的功能中表达出来；关于遗传信息的传递和翻译将在第二章中叙述。

结构

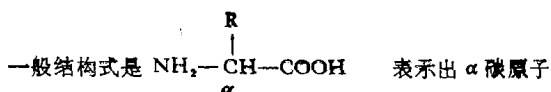
蛋白质功能的多样性反映在它的结构的多样性上。但是所有蛋白质都有两个共同的结构特征：它们只含有比例不一的20种不同类型的氨基酸(基本的氨基酸)，而且它们都是多肽，即通过肽键以有规则的方式连接起来的氨基酸聚合物。

虽然许多蛋白质含有300到400个氨基酸单位。但是它们的大小变化很大，可从50个氨基酸(如胰岛素)到2500个氨基酸(如肌球蛋白)，甚至到几百万分子量的巨大复合体。当蛋

白质以酸水解或以蛋白酶降解时，释出各种基本的氨基酸，而它们不必全都存在。偶尔在蛋白质中也发现一些非基本的氨基酸，但是这些非基本的氨基酸是在 20 种基本氨基酸生物合成蛋白质之后，才形成的 20 种基本氨基酸的衍生物。

图 1-2 中列出各种氨基酸的结构式，和常用的三个字母的简写。20 种基本氨基酸都是 α -氨基酸，但脯氨酸是例外，它是 α -亚氨基酸。脯氨酸在靠近羧基 ($-\text{COOH}$) 的 α -位置上有一个亚氨基 ($-\text{NH}$) 而不是一个氨基 ($-\text{NH}_2$)。 α -氨基酸的一般结构式是 $\text{RCH}(\text{NH}_2)\text{COOH}$ ，这个结构通式表明 α -氨基酸具有一个共同的结构部分，不同的只是 R 基团部分。图 1-2 表示这些不同的 R 基团，它们可以分为三大类。这三类不同的 R 基团使得氨基酸具有各种不同的性质，因而也使它们构成的蛋白质结构具有各种不同的性质。

第一类是非极性的或疏水的氨基酸，它包括那些具有脂族侧链的丙氨酸、亮氨酸、异亮氨酸、缬氨酸和脯氨酸，也包括具有芳香族侧链的苯丙氨酸和色氨酸，还包括含有硫的甲硫氨酸。它们的特征是对水没有亲和力，并且倾向于把 R 基团聚合起来以便排斥水。这对于蛋白质的折叠和脂蛋白结构的形成是一个重要的性质。第二类是不带电荷的极性氨基酸，它包括含羟基 ($-\text{OH}$) 的丝氨酸、苏氨酸和酪氨酸；含酰胺的天门冬酰胺、谷氨酰胺和含硫的半胱氨酸。为了方便起见，将甘氨酸也包括在这类中。甘氨酸是最小的，在性质上没有明显区别的氨基酸，同时它的 R 基团只是一个氢原子。这类氨基酸对水具有亲和力，在水中成为可溶的蛋白质。第



表示出 α 碳原子

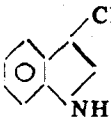
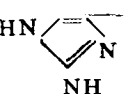
类别	名称	R基团	简写法
非极性或疏水的氨基酸	丙氨酸	CH_3-	Ala
	亮氨酸	$\text{CH}_3 > \text{CH}-\text{CH}_2-$ CH_3	Leu
脂族的氨基酸	异亮氨酸	$\text{CH}_3\text{CH}_2 > \text{CH}-$ CH_3	Ile
	缬氨酸	$\text{CH}_3 > \text{CH}-$ CH_3	Val
	脯氨酸	$\text{CH}_2 \begin{cases} \text{CH}_2-\text{到}\alpha-\text{N} \\ \text{CH}_2-\text{到}\alpha-\text{C} \end{cases}$	Pro
芳香族的氨基酸	苯丙氨酸	 - CH_2-	Phe
	色氨酸	 - CH_2-	Trp
含硫的氨基酸	甲硫氨酸	$\text{CH}_3-\text{S}-\text{CH}_2-\text{CH}_2-$	Met
不带电荷的极性氨基酸	甘氨酸	$\text{H}-$	Gly
含羟基(-OH)的氨基酸	丝氨酸	$\text{OH}-\text{CH}_2-$ OH	Ser
	苏氨酸	$\text{CH}_3-\text{CH}-$ OH	Thr
	酪氨酸	$\text{HO}-\langle \text{O} \rangle-\text{CH}_2-$	Tyr
含酰胺的氨基酸	天门冬酰胺	$\text{CONH}_2-\text{CH}_2-$	Asn
	谷氨酰胺	$\text{CONH}_2-\text{CH}_2-\text{CH}_2-$	Gln
含硫的氨基酸	半胱氨酸	$\text{SH}-\text{CH}_2-$	Cys
带电荷的极性氨基酸	天门冬氨酸	$\text{COO}^--\text{CH}_2-$	Asp
带负电荷的氨基酸	谷氨酸	$\text{COO}^--\text{CH}_2-\text{CH}_2-$	Glu
带正电荷的氨基酸	组氨酸	 - CH_2-	His
	精氨酸	$\text{NH}_3^+-\text{C}(=\text{NH})-\text{NH}-\text{CH}_2-\text{CH}_2-\text{CH}_2-$	Arg
	赖氨酸	$\text{NH}_3^+-\text{CH}_2-\text{CH}_2-\text{CH}_2-\text{CH}_2-$	Lys

图1-2 20种基本氨基酸的结构式和简写

三类是带电荷的极性氨基酸，它包括生理学上pH7左右的带负电荷的天门冬氨酸和谷氨酸，以及那些带正电荷的精氨酸、赖氨酸和组氨酸，尽管后者的电荷非常微弱。

旋光性是氨基酸的一种通性，这里只是简略地提一下。生物的分子通常是不对称的，并且有旋光性，即它们能旋转偏振光的平面。就氨基酸来说，除甘氨酸外，所有的氨基酸都有四种不同的化学基团，这些基团都连接在 α -碳原子上，从而构成不对称的分子，所以它们具有旋光性。幸而对于化学家的研究来说，在自然界中发现的19种氨基酸都具有同类的不对称性，因此它们的相反构型(对映体)在蛋白质中是不存在的，尽管有些相反构型确实在其他分子中存在。核酸的基本组件也有这种特性(参阅第1.5节)。已经弄清楚所有基本的氨基酸都是L-氨基酸，所以在本书讨论中都略去L-字首。

所有蛋白质都是多肽。它们是以肽键连接起来的许多氨基酸的线性排列，如图1-3所示。图中 R_1 和 R_2 表示属于不同氨基酸的R基团。这个线性排列或氨基酸顺序叫做蛋白质的一级结构。蛋白质的一级结构有两个特点是与我们了解的遗

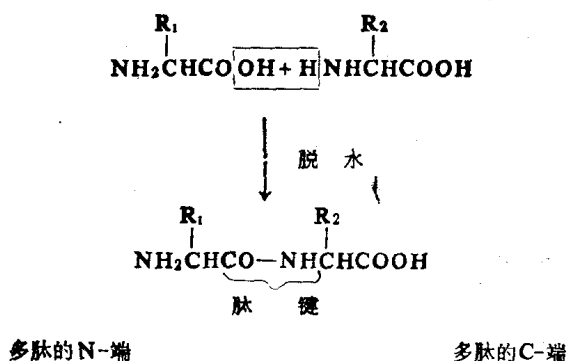


图1-3 肽键的形成