

目 录

第一章 绪论	(1)
第一节 分子生物学的含义	(1)
第二节 分子生物学发展简史	(2)
第三节 分子生物学在生命科学中的位置	(4)
第四节 分子生物学的现状和展望	(6)
第二章 生命的物质基础	(8)
第一节 生命是物质进化的产物	(8)
第二节 生命的分子逻辑	(9)
第三节 生物分子	(12)
第四节 生物分子的强键和弱键	(16)
第五节 生命系统和热力学	(20)
第六节 生物大分子的构象及其组装	(26)
第七节 蛋白质与其它生物分子的交互作用	(36)
第三章 细胞	(41)
第一节 引言	(41)
第二节 原核细胞与真核细胞的比较	(41)
第三节 原核细胞	(43)
第四节 大肠杆菌	(47)
第五节 真核细胞	(50)
第四章 DNA 的结构	(62)
第一节 DNA 结构的多样性和动态性	(62)
第二节 多核苷酸链	(62)
第三节 DNA 的双螺旋结构	(64)
第四节 A-DNA 和 B-DNA 结构的多态性	(67)
第五节 Z-DNA	(70)
第六节 DNA 的变性和复性	(72)
第七节 DNA 的形状、大小和序列组织	(75)
第八节 DNA 的精细结构	(78)
第九节 超螺旋 DNA	(83)
第十节 拓扑异构酶	(87)
第十一节 染色质的结构	(89)
第十二节 一些 DNA 序列的不寻常结构	(93)
第十三节 DNA 的限制图谱和序列分析	(97)
第五章 DNA 的复制和修复	(105)
第一节 DNA 复制概貌	(105)
第二节 DNA 的复制酶和相关蛋白	(109)

第三节	原核生物 DNA 的复制	(115)
第四节	真核生物 DNA 的复制	(124)
第五节	DNA 的修复	(129)
第六章	DNA 重组	(139)
第一节	引言	(139)
第二节	同源重组	(140)
第三节	位点专一性重组	(155)
第四节	转座重组	(159)
第七章	RNA 的合成	(174)
第一节	引言	(174)
第二节	原核生物 RNA 的结构及种类	(175)
第三节	RNA 酶促合成的特点	(177)
第四节	RNA 聚合酶	(180)
第五节	启动子	(184)
第六节	RNA 的酶促合成	(186)
第七节	RNA 转录后的加工	(190)
第八节	真核生物 RNA 的合成	(191)
第九节	RNA 的自我剪切与催化	(196)
第八章	遗传密码	(199)
第一节	遗传密码的破译	(199)
第二节	密码的简并性与变偶假说	(204)
第三节	密码子的使用频率	(206)
第四节	起始密码子与终止密码子	(208)
第五节	遗传密码的突变	(209)
第六节	遗传密码的改变	(215)
第七节	重叠基因与重叠密码	(217)
第九章	RNA 的翻译——蛋白质的生物合成	(220)
第一节	引言	(220)
第二节	核糖体及核糖体核糖核酸结构	(220)
第三节	氨基酸的激活与氨酰-tRNA 的合成	(233)
第四节	原核生物的蛋白质的生物合成	(241)
第五节	GTP 在蛋白质合成中的作用	(251)
第六节	真核生物的蛋白质的生物合成	(252)
第七节	蛋白质折叠与蛋白质生物合成中多肽链的修饰	(256)
第八节	蛋白质的易位与分泌	(262)
第十章	原核生物基因表达的调控	(271)
第一节	引言	(271)
第二节	细菌对营养的适应	(271)
第三节	组成蛋白与调节蛋白	(272)

第四节	操纵子学说	(273)
第五节	操纵基因的结构	(275)
第六节	乳糖操纵子	(278)
第七节	阿拉伯糖操纵子	(282)
第八节	氨基酸合成的操纵子	(284)
第九节	基因表达翻译水平的调节	(288)
第十一章	病毒	(290)
第一节	引言	(290)
第二节	病毒粒子的结构	(291)
第三节	病毒的基因组	(298)
第四节	病毒的复制	(317)
第五节	病毒基因的表达	(326)
第十二章	真核染色体和基因组的结构	(343)
第一节	真核染色体的结构	(343)
第二节	真核基因组的结构	(350)
第三节	细胞器中的基因组	(367)
第四节	mRNA 前体的剪接	(375)
第十三章	真核基因表达的调控	(386)
第一节	真核基因表达的调控	(386)
第二节	细胞周期和生长的调控	(437)
第三节	发育的基因调节	(444)
第十四章	分子生物学技术	(459)
第一节	引言	(459)
第二节	放射性同位素技术	(460)
第三节	核酸的提取和纯化	(462)
第四节	核酸的分子杂交和 DNA 的人工合成	(467)
第五节	目的基因的分离	(471)
第六节	核酸的序列分析	(476)
第七节	聚合酶链式反应技术及其应用	(483)
第八节	基因的表达和改造	(487)
第九节	RFLP 和 RAPD 及其应用	(492)
索引		(496)

第一章 绪 论

第一节 分子生物学的含义

生物学经历了一个漫长的研究历程。最早人们从研究动物和植物的形态、解剖和分类开始,以后进一步研究细胞学、遗传学、微生物学、生理学、生物化学,进入细胞水平的研究。到20世纪中叶以来,生物学以生物大分子为研究目标,分子生物学(molecular biology)开始形成了独立的学科,这是对生物界的认识不断深入的过程。

自从1838年Schleiden和Schwan证明动物和植物都是由细胞组成,Virchow(1858)提出细胞学说(the cell theory)之后,细胞学的研究得到迅速发展,随之遗传学原理也得到揭示。同时生理学和生物化学也随之兴起,以细胞为主要材料,进行深入地研究,使生物学的探索进入了细胞水平。

随着物理学和化学渗入生物学的研究领域,人们对细胞的化学组成的了解日益深化,我们对构成细胞的生物大分子,主要是蛋白质及核酸在生命科学中所起的作用有了深刻的认识。Sanger利用纸层析和纸电泳技术于1953年第一次揭示出胰岛素的一级结构,开创了以后数千种蛋白质序列分析的先声。不久Perutz及Kendrew(1953)利用X射线衍射技术解析了肌红蛋白(myoglobin)和血红蛋白(hemoglobin)的三维结构,使人们第一次能够洞察生物大分子的空间结构,从而了解这些蛋白质在运送分子氧中的特殊作用。这些研究结果使人们第一次能从分子水平了解生命物质的结构与功能。生物大分子结构与功能的研究乃成为生命科学中最重要的课题。

在核酸方面随着核酸化学研究的进展,人们揭示出核酸化学的许多规律之后,Watson和Crick(1953)共同提出了脱氧核糖核酸(DNA)的双螺旋模型(double helix model)。这个模型为揭开遗传信息的复制和转录的秘密铺平了道路。随后Crick又提出了中心法则(central dogma),明确了遗传信息传递的规律。从此以后,核酸的分子生物学得到了异乎寻常的迅速发展,使分子生物学成为生命科学中活力最强的学科。

从广义来讲,蛋白质及核酸等生物大分子结构和功能的研究都属于分子生物学的范畴,也就是从分子水平阐明生命现象和生物学规律。例如,蛋白质的结构、运动和功能,酶的作用机理和动力学,膜蛋白结构功能和跨膜运输等都属于分子生物学的研究内容。

不过目前人们通常采用狭义的概念,将分子生物学的范畴偏重于核酸(或基因)的分子生物学,主要研究基因或DNA的复制、转录、表达和调节控制等过程,当然其中也涉及与这些过程有关的蛋白质和酶的结构与功能的研究。本书也采用狭义的概念对基因分子生物学的基本原理进行讲述。至于专门问题如发育分子生物学、分子免疫学等,则请读者阅读有关专著。

第二节 分子生物学发展简史

随着化学及物理学的渗透,构成生物细胞的生物大分子的结构与功能的研究日益获得突破性的进展。早在1871年Miescher从死的白细胞核中分离出脱氧核糖核酸(DNA),迄今已有120年的历史。到了1928年Griffith发现肺炎链球菌(*Pneumococcus*)的无毒菌株与其被杀死的有毒菌株混合,即变成致病菌株。1944年Avery等人发现从致病力强的光滑型(S型)肺炎链球菌提取的DNA能使致病力弱的粗糙型(R型)转化成S型。如果加少量DNA酶,这种转化立即消失,但加入各种蛋白水解酶则不能改变这种转化。他们的实验充分证明引起细菌遗传改变的物质为DNA。

由于核酸化学的研究进展,Chargaff(1949)从不同来源DNA测定出4种核酸碱基,即胸腺嘧啶(thymidine, T)、胞嘧啶(cytocine, C)、腺嘌呤(adenine, A)和鸟嘌呤(guanine, G)。腺嘌呤与胸腺嘧啶的量和鸟嘌呤与胞嘧啶的量并不相等,即 $(A+T)/(G+C)$ 的比值随不同来源的DNA而有所不同。他发现鸟嘌呤的量与胞嘧啶的量总是相等,腺嘌呤与胸腺嘧啶的量相等,即 $G=C$, $A=T$,这个规律称为Chargaff规律。与此同时Wilkins及Franklin(1950~1952)用X射线衍射技术测定了DNA纤维的结构,它的衍射图像表明DNA具有典型的螺旋结构,并且由2条以上的多核苷酸链组成。当时,Pauling(1953)曾经提出DNA分子具有三股螺旋的设想(见Nature 1953, 171: 346~348)。几乎与其同时年青的Watson和Crick于1953年也在Nature杂志(171: 737~738)上提出了DNA双股螺旋模型。DNA以磷酸糖链形成了双股螺旋,脱氧核糖上的碱基按Chargaff规律构成双股磷酸糖链之间的碱基对。这个模型表明DNA具有自身互补的结构,根据碱基对原则,DNA中贮存的遗传信息可以精确地进行复制。他们的理论奠定了分子生物学的基础。

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DNA聚合酶I(DNA polymerase I),能使4种dNTP(即dATP, dGTP, dCTP和dTTP)连接成DNA。DNA的复制需要一个DNA作为模板。以后证明DNA的复制是一个非常复杂的过程,包含着许多种酶的参与。

DNA复制在分子生物学中是一个异常重要的问题。Meselson与Stahl(1958)用精彩的实验证明,DNA复制时DNA分子的两条链先行分开。他们用 ^{15}N 重同位素及密度梯度超速离心证明DNA复制是一种半保留复制。

Crick于1954年提出了遗传信息传递的规律,DNA是合成RNA的模板(template),RNA又是合成蛋白质的模板,称之为中心法则:



这个中心法则对以后分子生物学的发展起了极其重要的指导作用。

由于多年的研究编码组成蛋白质的氨基酸的遗传密码得到解决,对分子生物学的发展有重要推动作用。蛋白质由20种氨基酸组成,而DNA仅由4种核苷酸构成,按照中心法则,氨基酸与核苷酸的关系如何? Yanofsky和Brener(1961)提出了三联体(triplet)的设想,即3个

碱基编码一种氨基酸。这个问题经过 Nirenberg 和 Matthai(1963)的努力研究,编码氨基酸的遗传密码终于得到了破译。他们在无细胞系统中加入一定序列的人工合成的多核苷酸,即合成了一定序列的多肽链,充分证明编码 20 种氨基酸的遗传密码。

Khorana(1966)用实验证实了 Nirenberg 提出的遗传密码。Khorana 用有机化学方法合成了多聚脱氧核糖核苷酸,并以它为模板用 DNA 聚合酶 I 合成 DNA 链,然后他以 DNA 为模板用 RNA 聚合酶合成了 RNA 链,二者具有互补的关系。

蛋白质合成是分子生物学的重要课题,它的研究经历了很长的过程。早在 1953 年 Zamecnik 及其同事就开始在无细胞系统中利用放射性同位素标记的氨基酸研究蛋白质合成过程,发现蛋白质合成的场所为核糖体(ribosome)。他们还证明蛋白质合成需要 ATP 作为肽链形成的能源。氨基酸掺入蛋白质之前首先要与转移 RNA(tRNA)结合,它是由氨基酰合成酶(aminoacyl synthetase)催化的。在细胞总 RNA 中 tRNA 约占 10%,RNA 的 85%存在于核糖体(rRNA)中,1960 年以后利用 T4 噬菌体感染 *Escherichia coli* 作为系统,噬菌体侵染细菌后,寄主的 RNA 合成被中止,只有 T4 DNA 被转录成 T4 RNA。令人惊奇的是 T4 RNA 的碱基组成与 T4 DNA 非常相似,但它并不与 rRNA 结合形成核糖体。后来这种 RNA 携带 DNA 的信息转移到核糖体上合成蛋白质,故称为信使 RNA(messenger RNA, mRNA)。mRNA 约占总 RNA 的 4%。继 mRNA 被发现之后, Hurwitz, Stevans 及 Weiss 等人发现了 RNA 聚合酶,这种酶以 DNA 为模板利用 ATP, GTP, CTP, UTP 等合成 RNA,这就是转录(transcription)过程。

在细胞中蛋白质合成是受到控制的。例如, *E. coli* 中的 β -半乳糖苷酶(β -galactosidase)的含量就随着对它的需要而变化。当乳糖存在时它的含量就高,将乳糖分解成葡萄糖和半乳糖,而在乳糖不存在时,细菌合成的 β -半乳糖苷酶则极少。Monod 和 Jacob 于 50 年代末对此问题做了详细研究,提出了操纵子学说(operon theory),指出在操纵子中存在调节基因(regulatory gene),它可以产生阻遏蛋白(repressor),在乳糖不存在时阻遏蛋白就关闭结构基因,使之不能合成半乳糖苷酶。

DNA 是一个长链的生物高分子,在研究 DNA 重组、表达质粒的构造及它的碱基序列分析之前往往需要将 DNA 分子切割成为较短的片段,这就需要一种酶来完成。Smith 于 1970 年从 *E. coli* 中分离出第一个能切割 DNA 的酶。由于它能在 DNA 核苷酸序列的专一性位点上切割 DNA 分子,他将这种酶称为限制性酶(restriction enzyme)。以后很多种限制性酶陆续被分离出来,目前已有数百种限制性内切酶作为商品出售,给分子生物学研究带来极大的方便。

按照中心法则信息传递的方向是从 DNA 到 RNA,再从 RNA 到蛋白质。但是 RNA 在反转录病毒(retrovirus)中并非如此,它们能以 RNA 为模板合成单链的 DNA,然后再以这条 ssDNA(single-strand DNA)为模板合成互补 DNA(complementary DNA, cDNA)。以 RNA 为模板催化 DNA 合成的酶称为反转录酶(reverse transcriptase),它是由 Temin 和 Baltimore(1970)首次分别发现的。现在,我们可以利用反转录酶以分离得到的 mRNA 为模板合成 cDNA,从而进行基因结构及其表达的研究。

限制性内切酶的分离成功使得重组 DNA 成为可能。在此以前已经发现细菌中存在 DNA 连接酶(DNA ligase)。它能将被限制性酶切割的 DNA 片段连接在一起,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细胞之中,重组的 DNA 进行繁殖,于是产生了重组 DNA 的克隆(clone)。Berg 是重组 DNA 或基因工程技术

的创始人,于1980年获得了Nobel奖。

在研究分子生物学中了解DNA的核苷酸的排列顺序无疑是非常重要的。Sanger(1977)及Gilbert(1977)分别用与测定蛋白质序列截然不同的方法解决了DNA分子中碱基序列(DNA sequence)的复杂问题。Sanger采用的是酶法,而Glibert采用化学法。他们都将DNA分子中碱基的序列准确地测定出来,使我们对基因甚至基因组的结构得到了了解。

自Summer(1936)证明酶是蛋白质以来,已有50多年的历史,人们一直认为酶是蛋白质。但是近年有一个惊人的发现,即一些RNA也具有催化功能。Cech于1986年发现四膜虫(*Tetrahymena*)的核糖体RNA能够自我剪切。mRNA中的内含子能被RNA本身的催化作用准确无误地切除。这种内含子衍生出来的RNA在特定位点催化RNA链的剪切和连接,而其自身并不被消耗,完全符合酶的性质。这种催化剂被定名为核酶(ribozyme)。这一发现使人推测在生物进化的早期可能先形成RNA,然后以RNA为模板形成DNA。DNA后来代替RNA作为遗传物质,它的双螺旋结构比RNA单链更稳定,适宜于遗传物质的存贮,而RNA在核糖体中仍保留着催化性质。

由于分子生物学的研究对生命科学的发展起着巨大的推动作用,受到国际科学界的高度重视,许多位分子生物学家获得了诺贝尔化学奖或生理学奖。现在将历年来分子生物学家获得诺贝尔奖的情况简单介绍于表1中。

分子生物学从开始到如今只有50多年的发展历史,在人类文明史中只是短暂的一段,却使生物学发生了巨大变化,其进展可谓极其迅猛。由于无数分子生物学家的刻苦研究,使我们现在不但能从分子水平上了解DNA的结构、复制、转录和表达的详尽过程,而且对某些重要生物如果蝇(*Drosophila*)或拟南芥(*Arabidopsis*)复杂的发育过程有了深入地了解,使生物科学进入了一个新阶段,这在过去是办不到的。上面介绍的只是分子生物学的少数重要成就,我们不可能作全面详细介绍。关于分子生物学今后的展望,我们将在第四节加以讨论。

表1 分子生物学家获得诺贝尔奖一览表

人名	年代	获奖内容
Kornberg, A.	1959	DNA的复制
Watson, J. D.	1962	DNA双螺旋结构
Crick, F. H. C.		
Wilkins, M.		
Monod, J.	1965	操纵子学说
Jacob, F.		
Nirenberg, M.	1968	遗传密码的解析
Khorana, H. G.		
Holly, R.	1968	tRNA的结构
Temin, H.	1975	反转录酶的发现
Baltimore, D.		
Berg, P.	1980	DNA重组, DNA序列分析
Gilbert, W.		
Sanger, F.		
Cech, T.	1989	核酶(ribozyme)的发现

第三节 分子生物学在生命科学中的位置

分子生物学是从生物化学发展出来的一门学科。由于蛋白质及核酸生物化学的研究逐渐深入到这些生物大分子结构和功能的研究中,物理学家利用X射线衍射技术使蛋白质及核酸大分子的结构得到了解析,揭露出它们的三维结构,从而提出了脱氧核糖核酸(DNA)双螺旋模型。此后核酸的分子生物学得到了迅速发展,形成分子生物学一门独立的新学科。但是,分子生物学并不能与生物化学相分离,而是关系日益密切。例如,国际生物化学协会(The International Union of Biochemistry)现已改名为国际生物化学与分子生物学协会(The Interna-

tional Union of Biochemistry and Molecular Biology), 足以表明二者的关系十分密切。此外, 近年来蛋白质与 DNA 结合的研究日益受到重视, 成为晶体学(crystallography)的重要内容, 这个问题深入探讨将使分子生物学的重要机理, 如复制、转录、翻译、调控的本质得以充分阐明, 也就是蛋白质和酶究竟是如何使复杂的复制、转录及翻译过程进行的。DNA 虽然可以复制, 但是 DNA 本身却不能复制。RNA 可以转录, 但 RNA 本身却不能转录。复制和转录都是由蛋白质和酶实现的。翻译过程也是由核酸和蛋白质完成的。由此可见, 核酸和蛋白质的关系是相辅相成的。

目前分子生物学的研究虽然仍以 DNA 重组技术为主要手段, 但是愈来愈多的事实表明, 蛋白质和酶的研究在分子生物学研究中的重要地位, 例如, 研究基因的表达问题, 必然涉及到基因表达的产物, 也就是蛋白质和酶。而要想深刻阐明所表达的蛋白质和酶, 就必须彻底了解其结构和功能。因此, 在分子生物学的研究中, 蛋白质的纯化、一级结构、晶体的三维结构、溶液构象、光谱性质、酶的动力学等就成为必须研究的内容, 这样才能对基因表达的产物做出确切的解释。

分子生物学与微生物学的关系是密不可分的。早期分子生物学的研究对象都是原核生物, 特别是大肠杆菌(*E. coli*)。大肠杆菌的 DNA 和 RNA 及其复制、转录、翻译和调控过程我们已经了解得非常清楚。甚至我们现在用于分子生物学研究的各种质粒(plasmid)和限制性内切酶(restriction endonuclease)大都是 *E. coli* 中分离出来的。现在研究基因表达也往往是先用 *E. coli* 作为受体菌。难怪有人认为, 目前的分子生物学主要是“*E. coli* 的分子生物学”。当然, 目前分子生物学的研究对象早已不限于微生物, 动物和植物的分子生物学也已取得很大进展。由于微生物学与分子生物学的结合, 许多重要的微生物学课题已经达到了分子水平, 如生物固氮(biological nitrogen fixation)和植物微生物相互关系(plant-microbe interaction)的机理已日渐明朗。

遗传学无疑是与分子生物学关系异常密切的学科。分子生物学的发展极大地丰富了遗传学的内容, 产生了分子遗传学(molecular genetics)。分子遗传学与细胞遗传学、生化遗传学、数量遗传学、群体遗传学等, 构成遗传学的重要内容, 并且成为遗传学的重要部分。分子生物学极大地丰富了基因学说(gene theory), 不但大量基因的序列分析已测定清楚, 而且人类基因组(human genome)的结构也作为重大课题开展了研究。

细胞生物学与分子生物学也有密切的关系。现在, 细胞生物学已经不限于从细胞水平研究细胞的形态、结构与功能, 而进一步从分子水平探讨构成细胞各种组分的基因及其表达, 于是产生了分子细胞生物学(molecular cell biology)的新学科。这个学科正在蓬勃发展, 将会改变细胞生物学的面貌。

发育生物学研究动植物的生长发育, 是生命科学的重要内容。过去的组织学和胚胎学偏重于动植物结构和发育的形态描述。实验胚胎学虽然对生物的发育进行了细胞水平的研究, 但对发育的实质难以揭示。近年来分子生物学渗入到发育生物学以后, 发育生物学的面貌大为改观。现在我们已经认识到不论低等真核生物或高等生物, 它们的发育都是受基因组中 DNA 控制的, 并且按照特定的时间和空间顺序(temporal and spacial sequence)依次表达出来, 从而认识生物发育过程的实质。

目前生理学的研究也与分子生物学密切结合起来, 不论动物生理学或植物生理学都是如此。例如, 神经系统分子生物学的研究造成神经生物学的革命性变化。乙酰胆碱受体的分离纯

化和基因克隆已经完成。神经通道也已得到纯化和基因克隆。分子生物学渗入神经科学已产生了分子神经生物学。在植物生理学方面光合作用的研究已进入分子水平。一些植物的叶绿体DNA的全序列测定已经完成。光合作用有关的基因如光系统I, II, 电子传递链, 二磷酸核酮糖羧化酶/加氧酶(Rubisco)等基因的表达调控大量的研究都已完成了,使光合作用进入了一个新阶段。

不仅上述学科与分子生物学有密切关系,生物科学的许多学科也都进入分子生物学的研究领域,如分子免疫学、分子药理学、分子病理学、分子分类学等等,都从分子水平进行探索,使这些学科进入更高的阶段。

第四节 分子生物学的现状和展望

分子生物学的现状和发展前景是非常令人鼓舞的。90年代以来分子生物学在理论和技术方面都取得了重要进展,在DNA的复制、修复、转录、翻译和调控的分子机理等方面都得到了进一步的阐明,如拓扑异构酶I(topoisomerase I)的晶体结构、核糖体(ribosome)结构的研究使得我们对DNA的复制、翻译等的认识比过去更深入了。

分子生物学正在渗入到生物科学的有关学科之中,使得这些学科的面目大为改观。首先,在发育生物学方面,生活周期短的一些动植物如线虫(*Caenorhabditis elegans*)、果蝇(*Drosophila melanogaster*)、拟南芥(*Arabidopsis thaliana*)已成为发育生物学的重点研究对象,它们的发育过程很多已从分子水平得到了了解。例如,果蝇从卵发育成成虫的遗传信息是编写在基因组中的。一切性状都是由相关的基因所决定。在发育中基因的表达在时间和空间顺序上受到控制,如果蝇不能在幼虫时表达腿部的基因,只能在蛹期才能长出腿,也不能在蛹期产生卵,只能在雌性的成虫产生卵细胞。植物也是如此,它的生长发育过程也是在时间和空间上依靠各种基因依次表达的。这样,生物体的各种代谢过程才得以有条不紊地进行。例如,植物开花的基因十分复杂,花的各种器官由许多基因控制着,花瓣的颜色由多种次生物质代谢产

进展极快,在全世界范围业已达到平均每天能解析出三种蛋白质晶体结构的速度,而30年前,10年间只解析出一种蛋白质结构,结构生物学这样的高速度发展对生物科学做出了重大贡献。高分辨的蛋白质晶体结构使我们清晰地看到蛋白质分子中多肽链如何折叠成 α 螺旋、 β 折叠片、 γ 转角以及整个分子的折叠情况,甚至蛋白质(酶)与配基(ligand)的结合情况。在结构生物学中DNA与蛋白质的相互作用是一个重要领域,它对分子生物学的理论研究至关重要。例如,在核酸结构中除去双股DNA(DNA duplex)外,还发现了三股DNA(DNA triplex)和四股DNA(DNA tetraplex),三股DNA在抑制真核基因表达中有重要作用,而四股DNA存在于染色体的端粒(telomere)中,有稳定染色体结构的作用。目前,晶体学家已经解析出大量与DNA相互作用的蛋白质,如各种限制性内切酶(*EcoRI*, *BamHI*),各种阻遏蛋白(repressors, CAP, trp, Cro, 434, Arc, MetJ)、DNA修复蛋白(DNA repair protein)、TATA box蛋白、组蛋白(histone)、转座酶(transposase)等。拓扑异构酶I(topoisomerase I)的三维结构也已解析出来,它是一个很特殊的酶,具有环状结构,它能在瞬间打断一股DNA,将另一股DNA穿过断裂处,然后又封闭裂口,使两股DNA重新组合。拓扑异构酶I在DNA复制、转录及重组中起重要作用。此外,分子生物学对结构生物学的发展也有很大帮助。从动植物、微生物细胞中分离纯化蛋白质是一项十分艰苦的工作,由于大多数蛋白质在细胞中的含量都很低,要想分离纯化出某种蛋白质,进行晶体培养,是相当困难的。现在,我们可以将编码某种蛋白质的基因用分子生物学技术进行克隆,用聚合酶链式反应(PCR)扩增,然后与融合蛋白质粒(fusion protein plasmid)构建成表达质粒,用它转化大肠杆菌(*E. coli*)。然后进行发酵培养,即可得到大量含有该种蛋白质的菌体。从这些菌体中即可分离纯化出足够数量的该种蛋白质,从而培养晶体就容易得多了。

蛋白质工程(protein engineer)技术与分子生物学结合是促进分子生物学发展的一条途径,采用定点突变(site-directed mutagenesis)方法使基因结构发生改变,从而可以改变基因表达产物中的氨基酸残基,就有可能使我们了解蛋白质中每个氨基酸甚至每个化学基团所起的作用。

目前,动植物基因组(genome)的研究已成为分子生物学的重大课题。最近线虫(*Caenorhabditis elegans*)基因组即将测定完成,它由1000万个核苷酸组成。我国科学家正在进行水稻(*Oryza sativa*)基因组的DNA序列测定,这项工作的完成必将对植物分子生物学及农业科学做出重大贡献。目前,人类基因组(Human genome)的研究是国际上的重大科学课题,世界上许多国家投入大批人力、物力研究人类染色体DNA的全序列。这项宏伟工程的完成将会对分子生物学做出巨大贡献,并将对医学特别是遗传病的诊断治疗做出贡献。

今后,分子生物学在实际方面的发展也是令人乐观的。在医学方面,基因治疗(gene therapy)目前已在研究之中。在动植物方面转基因动植物的培育,也已进入重点研究之列,成为生物技术(biotechnology)的主要内容。将苏云金杆菌毒蛋白(*Bt*)基因转入棉花已获得抗棉铃虫的抗虫棉,并已取得经济效益。利用反义RNA(antisense RNA)技术延长果实、蔬菜的保鲜期和改变花卉的花色也已取得成果,这些都是分子生物学在实际方面的贡献。随着分子生物学和生物技术的发展,它们在医学和农业方面必将对人类的健康和生活做出更大的贡献。

(阎隆飞)

第二章 生命的物质基础

对生命现象本质的认识,历来存在两种截然不同的看法:生命是一种超自然现象和生命是物质的一种属性。随着物理学和化学对生物学的渗透和融合,作为生物学语言的生物化学出现了。而蛋白质和核酸等生物大分子结构和功能关系的阐明进一步使人们得以从分子水平上来认识生命现象的总体。事实表明,生物机体中物质的转化以及伴随的能量变化,各种化学组分在时空上的相互关系,以及驾驭这一切的遗传信息的传递无一不遵循物理学和化学的基本法则。这些都证明生命不是超越自然的,它的本质即是物质运动的一种状态。

第一节 生命是物质进化的产物

洪荒时代的地球是一个无生命的世界,它笼罩在一种还原性大气中,当时大气的成分主要是 H_2 , CH_4 , NH_3 , N_2 , CO , CO_2 , H_2S 等。以后随着水分由地球内部外溢,海洋形成了。从此,地球上的化学反应活跃了起来,地球上的物质开始了化学进化。无机化合物藉地热、放电、紫外线、宇宙射线等的能量合成了简单的有机化合物,随后氨基酸、糖、脂肪酸、核苷酸、卟啉化合物等相继形成。地球上出现了非生物合成的有机化合物。随着有机化合物的蓄积,原始生命终于在某种特定环境中出现了,这大约发生在 30 亿年前,即前寒武纪年代,物质进化从此步入生物进化(图 2-1)。最原始的生物是一些异养厌氧生物,它们以环境中的非生物合成有机成分为养料,

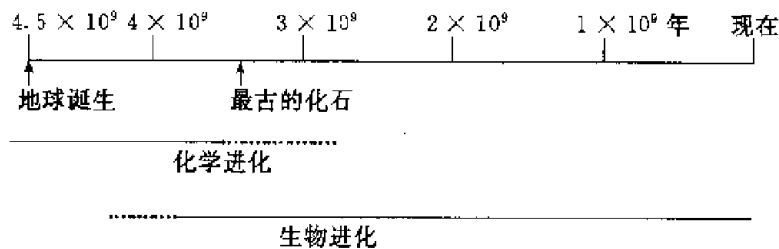


图 2-1 地球上的物质演化进程

通过酵解方式获取生物可以利用的能量形式——ATP。但环境中的有机成分有限,这就限制了原始生物的发展。随着时间的推移,一种光合自养生物出现了,它使还原性大气中出现了氧。氧虽不是生命所必需的,却是高水平的能量代谢所不可少的。氧对生物进化孕育着巨大的潜力。有氧生活虽更能满足高等生物所需的能量,不过氧对原始厌氧生物却具有毒性。大气中氧的出现推动了生物进化,原始厌氧生物终于获得了对氧毒性的防护机制。糖的酵解发展为有氧分解,原核生物发展为真核生物。这些变化为生物进化在遗传上和代谢上准备了条件,生物进化加速了。

生命,不管是其低级或高级形态都显示出与无生命物体明显不同的特点。生命的出现是物质进化的一个转折点。

生命所具有的特征是:

1. 生物体是由许多称为生物分子(biomolecule)的有机化合物组成的。活机体中,生物分子是有组织的,它们按一定的层次形成一高度有序的系统。活机体能维持这一有序性直至死亡。无生命物质则相反,它们总是自发地向无序方向发展。

2. 活机体不断地与环境发生物质和能量的交换。活机体中存在着一个具有一定顺序、相互协调、可自我调节的代谢网络,其中各个代谢反应都由相应的酶催化。细胞和机体与环境保持在一个远离平衡态的稳态(steady state)中。与此相反,无生命物质总是趋向于与环境达成平衡。

3. 生物体能够精确地自我复制,即机体的形状、结构和功能等能够代代相传,能够生长、繁殖。与此相反,无生命物质则不能自我生长和繁殖。

什么是生命是一个难以简单界定的问题,它的答案还得从活机体的自身去寻找。

第二节 生命的分子逻辑

活细胞或机体是由许多无生命的有机化合物组成的。但细胞不是一般意义上的口袋,会包容所有相遇的物质,而是有选择的摄取。在物质进化过程中,只是特定生物分子的集聚才导致生命的出现。在物质化学进化的末期,核苷酸、氨基酸、脂肪酸、叶啉等小分子化合物出现了。根据现有资料,人们认为这时可能出现了关键性的一步,即是核苷酸缩聚成了RNA。它是第一个具催化功能的基团,它能催化自我复制。RNA自我复制中产生的变异体引起了氨基酸的缩聚,蛋白质出现了。以后,DNA替代RNA成为遗传信息的载体,脂类物质形成的半通透性膜使前述的分子聚集体形成为细胞。

上述生命发生过程虽在一定程度上仍是一种假设,犹待进一步验证。不过,这一假设已指明,在细胞和机体中,生物分子是高度组织化的,形成了严密的系统,生命现象是它的总体表现。

一、机体化学组成上的统一性和多样性

活机体,不论是原核或是真核生物都是由核酸、蛋白质、脂质和糖类等生物大分子和一些小分子化合物及无机盐组成的。表2-1列出的是原核生物大肠杆菌的分子组成。事实上,所有生物的化学组成都可以分为这几类,

这就是其统一性之所在。不同生物在化学组成上的差异不在于分子的类别,而在于每一类分子的质与量。如以大肠杆菌与最简单的真核生物酿酒酵母相比,前者含1个染色体,其DNA含 4.7×10^6 bp(base pair, bp),而后者含16个染色体,其DNA含 1.4×10^7 bp。不仅如此,DNA的核苷酸序列也不同。DNA质和量上的不同决定了细胞所含的RNA的不同,主要是mRNA的不同,从而也使所含的包括酶在内的蛋白质的质和量不同,转而影响机体的代谢和别的化学组成。机体这种代谢上的差异就反映为生物性状上的多样性。生物的多样性是以其统一性为基础的。

表 2-1 大肠杆菌细胞的分子组成

	占总重的百分数	分子的种数
水	70	1
蛋白质	15	3 000
核酸		
DNA	1	1
RNA	6	>3 000
多糖	3	5
脂质	2	20
生物大分子的结构		
单位和中间体	2	500
无机离子	1	20

二、生物大分子的高聚物特性

核酸(DNA 和 RNA)、蛋白质、多糖和脂质是组成生物体的 4 类生物大分子。DNA 是生物体中遗传信息的原初载体。DNA 通过复制使遗传信息由亲代流向子代,通过转录使特定基因的遗传信息转换成相应的指令——mRNA,后者指导氨基酸按一定的顺序连接成特定的多肽,然后折叠成相应的蛋白质。蛋白质是遗传信息的体现者。核酸和蛋白质合成代表生命活动中遗传信息流动的主线,它驾驭着生命活动的进行。核酸和蛋白质的高聚物特性正是实现这种信息流动的基础,核酸分子的骨架是由核苷酸通过 3',5'-磷酸二酯键连接成的多核苷酸链,核苷酸是其单体。构成 DNA 和 RNA 的分别是 4 种脱氧核糖核苷酸和核糖核苷酸。不同的核糖核苷酸(和脱氧核糖核苷酸)的区别在于其碱基的差异。蛋白质分子的骨架是由 20 种氨基酸通过肽键连接成的多肽链。20 种氨基酸的区别在于其侧链(R 基)的差异(图 2-2)。这就极大地简化了遗传信息的转化,使其成为 4 种核苷酸和 20 种 α -氨基酸连接顺序间的转换,亦即核酸语言转换成了蛋白质语言。在转录中,DNA 的碱基顺序决定了新合成的 mRNA 的碱基顺序,这是遗传指令的发送。在翻译中,mRNA 上的碱基顺序规定了新合成的多肽链的氨基酸顺序,而氨基酸侧链的结构和性质则决定了多肽链可折叠成的稳定构象和形成相应的功能。这是指令转换为功能的过程。

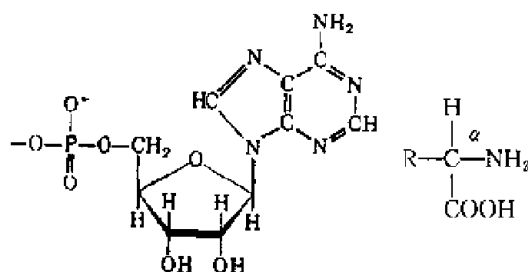


图 2-2 核苷酸和 α -氨基酸

除了核酸和蛋白质外,多糖和脂质是组成细胞的另两类重要的物质。多糖的单体是单糖。根据单糖的种类,多可分为同多糖(homopolysaccharide)和杂多糖(heteropolysaccharide)。同多糖由一种单糖组成,杂多糖由一种以上的单糖组成。淀粉、糖原和纤维素属同多糖,都由 D-葡萄糖组成,但它们的分子内糖苷键的类型有所不同。糖蛋白和糖脂的糖链属杂多糖,它们不仅组成单糖的种类多,并且单糖残基间糖苷键的类型也多样化。这些寡糖链是机体中分子的识别标志。寡糖链的组成和结构也是受遗传信息控制的,它们是在特定酶的催化下形成的。

脂质是一个松散的概念。它包括许多化学上不同的物质,如油脂、磷脂、类固醇、类胡萝卜素和脂溶性维生素等。它们的共同特点是难溶于水。脂质从结构上考虑不属于高聚物,但脂质中的磷脂等是两亲分子(amphipathic molecule),即分子的一部分亲水另一部分亲脂,在水中能藉疏水交相互作用聚集成双分子层,是膜的基本组成物质之一。磷脂分子中脂肪酸链的不同会影响它们在膜中排列的规整性,从而影响膜的一些特性,如流动性等。这类似于单体对高聚物的影响。

从以上的讨论,我们不难看出表 2-1 所列的 6 000~7 000 种分子是由 4 种核糖核苷酸,4 种脱氧核糖核苷酸,20 种 α -氨基酸以及少数几种单糖和脂肪酸等组成的,这些生物大分子又进而形成超分子复合物和细胞(图 2-3)。

细胞	
超分子复合物	染色体,核糖体,膜,微管等
生物大分子	DNA,RNA,蛋白质,多糖等
单体	核苷酸,氨基酸,单糖,脂肪酸等

图 2-3 细胞的结构层次

三、生物大分子单体在代谢中的地位

各种单体不仅是生物大分子的结构单位,又是合成机体的许多重要组分的前体(图 2-4)。

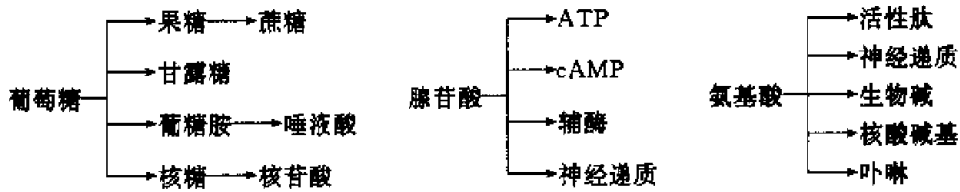


图 2-4 生物大分子单体是一些重要生物分子的前体

生物大分子的单体,尤其是单糖、脂肪酸以及氨基酸也是机体中间代谢的重要环节,它们循不同而又相互连接的途径逐步降解并汇集到少数共同产物,如 α -酮戊二酸(C_5),草酰乙酸(C_4),丙酮酸(C_3)和乙酰辅酶 A(C_2),然后进入三羧酸循环并通过呼吸链生成二氧化碳(C_1)和水(图 2-5)。这是一放能过程,所释放出的能转化为代谢中的能量通货——ATP。分解代谢和合成代谢是新陈代谢的两个基本方面。从总体说,分解代谢是放能过程,合成代谢是需能过程。分解代谢释放出的自由能使 ADP 和无机磷酸(P_i)形成 ATP,合成代谢所需的能可由 ATP 裂解为 ADP 和 P_i 来提供。ATP 是这两大过程间能量转移的基本环节(图 2-6)。

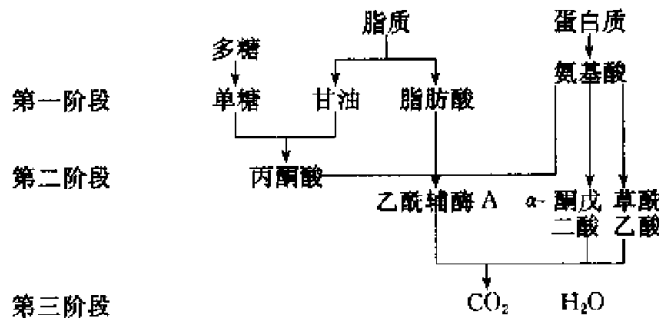


图 2-5 多糖、蛋白质和脂质的降解过程

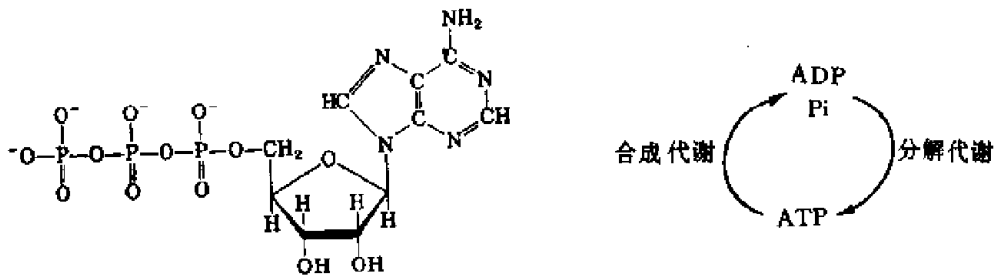


图 2-6 合成代谢和分解代谢间的能量转移

四、活机体是一远离平衡态的开放系统

分解代谢和合成代谢间的能量转移并不说明两者间存在平衡,而只是说机体中 ATP 处在稳态中。在活机体中,ATP 转化为 ADP 和 Pi 的速率导致 ADP 和 Pi 重新合成 ATP,故 ATP 的总浓度趋向于不变,但细胞和机体是远离平衡态的。

细胞的各部分是不均匀的,各类生物分子的分布存在着差异,各部分间也存在着分子和离子的浓度梯度以及电位差等,正是这种差别使细胞各部分间存在物质和能量的流动。同样,细胞、机体与环境间也存在着差异,这种差异较细胞内各部分间的差异更为巨大。一膜之隔,外部是无生命的,死气沉沉;而内部是有生命的,生机盎然。内外一旦平衡,死亡立即降临。

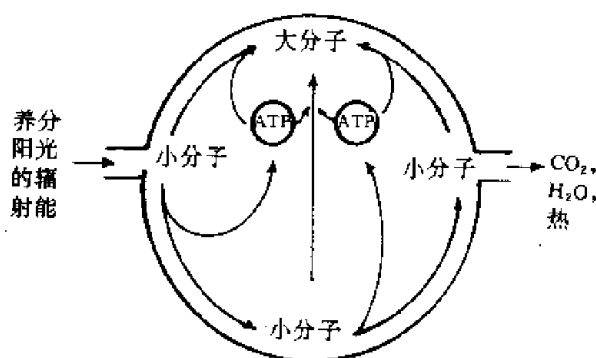


图 2-7 细胞为一开放系统

已如上述,活机体是一高度有序的系统,这是生命的基础所在。机体为了保持自身的这种有序性就必须从环境吸收淀粉、脂肪和蛋白质等有机养分,再通过分解代谢取得自由能,也可从阳光摄取辐射能。在此同时,机体向环境排出 CO_2 , H_2O 等简单分子以及在等温、等压条件(这是机体的工作环境)下难于利用的热。细胞为一开放系统(图 2-7)。

细胞的代谢构成了一个复杂的网络,各个代谢反应都是酶促反应。代谢反应基本上都是可逆的,但在活细胞中,由于反应产生的产物很快被移去,故反应难以达成平衡。这使细胞中的代谢物保持一定流向。远离平衡态是活机体的一个基本特征,细胞是一个远离平衡态的开放系统。

第三节 生物分子

生命是物质进化到达一定阶段后的产物。在此之前存在于地球上的矿物质并不能导致生命的出现,只有当核苷酸、氨基酸、脂肪酸、单糖等形成后,生命的诞生才具备了条件。

一、生命的化学元素

生命的产生具有明显的化学背景。在多于 90 种的天然存在元素中,只有约 30 种是机体必需的,其中碳、氢、氧、氮占多数细胞质量的 99% 以上。假如以地壳、海水和人体的元素组成做成一比较(表 2-2),我们不难发现人体的元素组成相对更接近于海水。这和构成细胞的有机化合物形成于陆地,而生命则发生于海水的假设相一致。

水是机体中一种大量存在的化合物,它占机体重量的 70% 或更多。许多证据表明,生命发生于水中。水影响着细胞中生物分子间相互作用,水的电离特性也影响着生物分子的功能,水也使细胞质呈现出一定的结构性。水对生命现象具有很大的重要性。除了氢、氧外,碳构成了生物分子的骨架。碳具有一个非常独特的性质,即碳原子能连接成各种稳定的链或环,所以,生物机体不选择与碳同族的、在地壳中含量丰富的硅。生物分子除碳以外还含氢、氧、氮、磷、硫,

这 6 种元素有一个共同的特点,即都是轻元素,都能形成稳定共价键,除氢外还都能形成重键(双键和三键)。它们形成了具不同结构和性能、适合表现出生命现象的化合物。

表 2-2 地壳、海水和人体组成元素的丰度

	海水(%)		人体(%)		地壳(%)
H	66	H	63	O	47
O	33	O	25.5	Si	28
Cl	0.33	C	9.5	Al	7.9
Na	0.28	N	1.4	Fe	4.5
Mg	0.033	Ca	0.31	Ca	3.5
S	0.017	P	0.22	Na	2.5
Ca	0.0062	Cl	0.08	K	2.5
K	0.0060	K	0.06	Mg	2.2
C	0.0014	S	0.05	H	0.22
		Na	0.05	C	0.19
其它	<0.1	其它	<0.01	其它	<0.1

除了上述 6 种元素外,细胞还含有许多微量元素:铁、铜、锰、锌、钴、钼、硒、钒、镍、碘、镁等。

这些元素的含量虽甚微,但它们对生命活动却具有十分重要的功能。其中许多是酶的不可缺少的组分,有的是氧化还原中的电子载体,有的参与代谢的调节,有的是呼吸作用和光合作用所不可缺少的,等等。

二、生物分子是碳的化合物

生物分子是碳的化合物,碳构成了它们的骨架,在生物分子中,碳原子可通过单键相连接,也可以通过双键(生物分子很少含三键)相连接;碳原子可以连接成链,也可以连接成环。分子中,饱和碳原子的 4 个价键分布于空间,故呈立体型;以双键连接的碳原子上的 4 个价键分布于同一平面中,故呈平面型(图 2-8)。这是生物分子的结构呈现出多样性的基础。

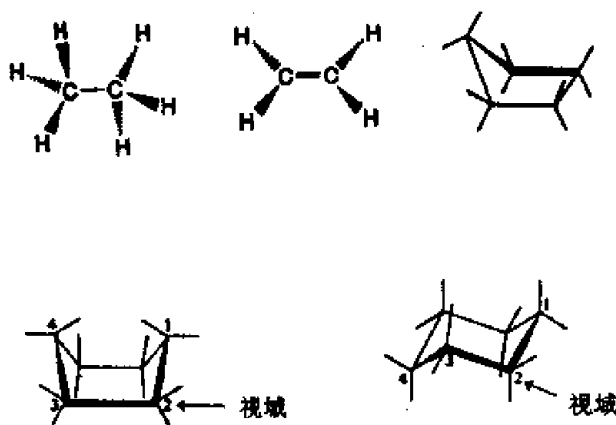


图 2-8 有机化合物的分子模型

碳氢化合物是有机化合物的母体,化学上稳定,但当它们的分子中引入一定的官能团后,化学活动性就相应增加了。生物分子中常见的官能团有羟基、羰基、羧基、氨基、巯基等(图 2-9)。

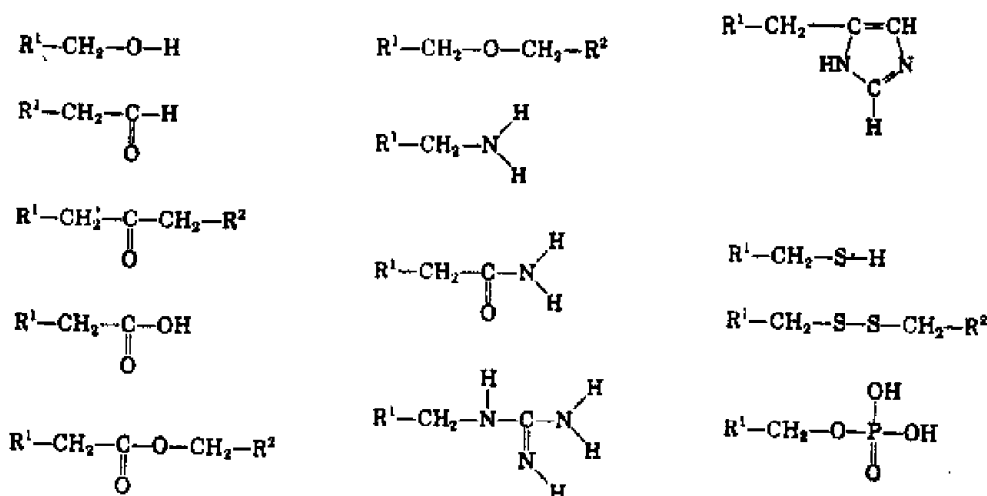


图 2-9 一些常见的官能团

生物分子除碳环外还常含杂环(由碳及非碳原子构成的环状结构),它们在生命活动中起着重要作用,如核酸的嘧啶和嘌呤碱基,氨基酸侧链中的咪唑和吡咯基以及血红素和叶绿素中的卟啉等。杂环化合物常含共轭双键,故呈扁平型。

生物分子的各种骨架以及它们具有的不同官能团使它们在结构和功能上显示出高度的多样性,从而满足生命活动的需要。

三、生物分子的构象和构型

关于一个化合物的分子结构,我们不仅要知其构成原子的种类和数目,相互间的连接方式和连接顺序,还要确定原子在空间的排布。分子中组成原子的空间排布对生物分子,尤其是生物大分子十分重要,因为三维的细胞是由三维的生物大分子组成的,而生物分子间高度专一的相互作用也是通过空间关系实现的。构象和构型就是指分子中组成原子在空间的相互关系。

1. 构象 乙烷分子含 2 个碳和 6 个氢原子,这 6 个氢原子在空间的关系不是固定的,这是由于两个碳原子可以单键为轴心而相对转动。不过两个碳原子的相对转动不是完全自由的,这是由于当碳原子作相对转动时,非键合的氢原子间会发生不同程度的干扰,从而使分子的势能发生改变(图 2-10)。图中表示的只是乙烷分子中氢原子空间排布的两种极端方式,一种称为重叠式;另一种称为交叉式。重叠式的势能最高,交叉式最低,其它方式则介于两者之间。如乙烷那样,由于有关原子相对转动所造成的相应原子或基团在空间的不同排布称为构象(conformation)。

乙烷分子只含两个碳原子,它的构象变化较为简单。但当乙烷分子的两个碳原子各有一个氢为甲基取代时(即形成丁烷),情况就复杂得多。如仅取第二和第三两个碳原子,它就可以形成如图 2-11 的几种主要构象,如再考虑第一和第二及第三和第四碳原子的相对转动,则不仅构象变化更为复杂,并且各个极端构象间的势能差也将扩大。不同构象能量上的差异来自有关原子或基团的大小和性质,不同大小的原子或基团会产生不同的空间障碍,而基团的性质如是否具有极性会产生排斥或吸引。表 2-3 列出了一些原子的大小,其中共价半径代表成键时该原子的半径,而范德华半径则代表从一个原子的中心到它所接触的非键合原子的距离。共价半径之和即为该两原子间的键长,而范德华半径则决定分子可达到的紧密程度。