

21 世纪

周虹 主编
徐文弟

供研究生用
高等医学院校教材

生物化学与 分子生物学 高级教程

 科学出版社

171

05-43
Z74

21世纪高等医学院校教材

(供研究生用)

生物化学与分子生物学 高级教程

周虹 徐文弟 主编

科学出版社

2002

内 容 简 介

本书是根据研究生与七年制医学生课程教学要求而编写的。全书主要分为两部分。第一部分讲述生物化学与分子生物学基础理论,包括生物大分子的结构和功能、基因表达调控、生物膜、信息传递、代谢调节、生物氧化等章节;第二部分叙述相关医用生物化学内容,如血液、神经生化、肿瘤等。考虑到细胞对生命现象的重要意义和生命科学研究技术的发展,增加了细胞分化和生物技术等内容。本书内容丰富,实用性强,可供全国医学院校各专业研究生及七年制医学生使用,也可供有关人员参考。

图书在版编目(CIP)数据

生物化学与分子生物学高级教程/周虹,徐文弟主编.-北京:科学出版社,2002.1

21世纪高等医学院校教材.供研究生用

ISBN 7-03-009709-2

I. 生… II. ①周…②徐… III. ①生物化学-研究生-教材②分子生物学-研究生-教材 IV. ①Q5②Q7

中国版本图书馆CIP数据核字(2001)第055589号

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>

新蕾印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2002年1月第一版 开本:850×1168 1/16

2002年1月第一次印刷 印张:22 1/2

印数:1-5 000 字数:466 000

定价:35.00元

(如有印装质量问题,我社负责调换〈环伟〉)

《生物化学与分子生物学高级教程》

编写人员

主 编 周 虹 徐文弟

副主编 于晓光 刘兴汉

编 者 (按姓氏笔画排序)

于 泓 于晓光 王秀宏 王爱民

尹 琄 刘兴汉 初彦辉 李 晖

张 涛 张雪峰 范业鹏 杨春波

欧 芹 周宏博 周 虹 赵炜明

高 旭 郭新民 徐文弟 曹 军

前 言

生物化学与分子生物学是医学院校重要的基础课程,为阐述疾病发生机制和现代医疗技术原理奠定了基础。

根据研究生与七年制医学生生物化学与分子生物学课程教学要求,我们与几所兄弟院校编写了这部教材,供硕士研究生和七年制学生基础教学使用。

本教程主要分为两部分,第一部分讲述生物化学与分子生物学基础理论,包括生物大分子的结构与功能、基因组、基因表达调控、生物膜、信息传递、酶、代谢调节、生物氧化等章节;第二部分叙述相关医用生物化学内容,如血液、神经生化、肿瘤等。考虑到细胞对生命现象的重要意义和生命科学研究技术的发展,增加了细胞分化和生物技术等内容。

作为研究生用的生物化学与分子生物学教材,我们是初次编写。在内容的选择、材料的取舍、章节的安排等方面一定存在着许多不足和问题,希望广大同行多提宝贵意见。

在编写过程中,王均衡教授、张乃忠教授对书稿进行了审阅,并提出了许多宝贵意见,同时还得到了哈尔滨医科大学及基础医学院有关部门及领导的支持和帮助,在此一并表示感谢。

编者

2001年7月

目 录

前言

第一章 生物大分子的结构与功能	(1)
第一节 蛋白质的结构与功能.....	(2)
第二节 核酸的结构与功能.....	(26)
第三节 糖复合物的结构与功能.....	(51)
第二章 基因组	(80)
第一节 不同生物基因组特点.....	(80)
第二节 人类基因组.....	(86)
第三节 人类基因组研究.....	(88)
第四节 蛋白质组学.....	(98)
第五节 基因组研究与生命科学工业.....	(107)
第三章 基因表达与调控	(113)
第一节 概述.....	(113)
第二节 原核基因的表达调控.....	(114)
第三节 真核基因的表达调控.....	(119)
第四章 生物膜	(129)
第一节 细胞的膜系统.....	(129)
第二节 生物膜的结构.....	(131)
第三节 生物膜的功能.....	(137)
第五章 信息传递与受体	(141)
第一节 信息分子.....	(141)
第二节 受体.....	(143)
第三节 生物信息传递机制.....	(151)
第四节 信息传递障碍及受体病.....	(157)
第六章 酶	(161)
第一节 酶的分子结构.....	(161)
第二节 酶的催化机制.....	(166)
第三节 酶促反应动力学.....	(176)
第四节 酶活性调节.....	(193)
第五节 酶与医学的关系.....	(205)

第七章 代谢调节	(207)
第一节 代谢调节的重要意义.....	(207)
第二节 细胞水平的代谢调节.....	(212)
第三节 激素.....	(219)
第四节 整体水平的调节.....	(227)
第五节 研究代谢的方法.....	(231)
第八章 生物氧化	(235)
第一节 生物氧化的基本概念.....	(235)
第二节 呼吸链.....	(236)
第三节 氧化磷酸化作用.....	(240)
第九章 血液	(249)
第一节 血液的化学成分.....	(249)
第二节 血浆蛋白质.....	(251)
第三节 血浆脂蛋白.....	(253)
第四节 血红蛋白.....	(262)
第十章 神经生化	(270)
第一节 神经元膜和突触的分子结构特点.....	(270)
第二节 神经传导的分子基础.....	(275)
第三节 中枢神经递质.....	(282)
第四节 神经多肽.....	(286)
第十一章 细胞分化	(292)
第一节 细胞分化的基本概念.....	(292)
第二节 细胞的发育.....	(294)
第三节 细胞质在细胞分化中的决定作用.....	(297)
第十二章 肿瘤	(300)
第一节 肿瘤发生的分子生物学基础.....	(300)
第二节 肿瘤转移的分子基础.....	(313)
第三节 肿瘤的基因治疗.....	(316)
第十三章 生物技术	(321)
第一节 生物技术概论.....	(321)
第二节 基因工程.....	(325)
第三节 分子杂交、聚合酶链反应、DNA 序列分析和限制性片段长度多态性分析	(340)
第四节 转基因动物和克隆动物.....	(347)

第一章

生物大分子的结构与功能

生物分子的相对分子质量范围一般在 $18(\text{H}_2\text{O}) \sim 10^{11}$ (如真核 DNA 分子), 通常可将所有的生物分子简单地分为两类: 一类是小分子 (small molecule), 其相对分子质量一般小于 500, 为简单的单体物质; 一类是大分子 (macromolecule), 其相对分子质量一般在 2 000 以上, 通常要大得多, 一般是聚合物。机体内重要的生物大分子有蛋白质、核酸和多糖, 它们是生命体结构和功能的重要基础。

研究生物大分子特定的空间结构及结构的运动变化与其生物学功能关系的科学被称为结构生物学 (structural molecular biology)。众所周知, 分子生物学已成为生命科学和医学中的领先学科、核心学科和共同语言。近 20 年来, 几乎每年的医学和生理学诺贝尔奖以及一些化学奖都授予了从事生物化学和分子生物学的科学家, 这个事实本身就说明了分子生物学在整个生命科学中的战略地位和重要意义。而生物大分子的结构和功能恰恰反映了分子生物学的核心问题。结构生物学已成为生物学各前沿领域的基石。虽然结构生物学名称的提出已有近 30 年, 但结构生物学时代的真正开始不过是近几年的事。如果说 20 世纪 50 年代 DNA 双螺旋结构的发现使生物大分子研究的主流从蛋白质转向核酸, 那么现在可以说又从核酸转回蛋白质了。结构生物学尤其依赖于新技术和新仪器的应用。核酸和蛋白质的提取、合成和顺序测定已经可用自动化仪器来完成; DNA 重组技术在蛋白质结构和功能的研究中日益发挥重要作用; 采用标记和酶逐步降解技术, 使多糖结构得以阐明; X 射线晶体学分析可提供大分子结构的高清晰度图像; 质谱和各种光谱技术的进一步改善, 可以提供更多的结构信息。结构研究的发展方向, 一是趋向更大、更复杂的结构, 从研究大分子的结构进而研究超分子复合体, 如病毒颗粒、细胞器结构和膜复合体, 使结构研究从大分子到细胞得以衔接; 二是由静态发展到动态。目前, 已能在毫秒数量水平上测定蛋白质的动态活动, 如蛋白质变性和新生肽链的折叠情况, 酶与底物作用时构象变化等。

第一节 蛋白质的结构与功能

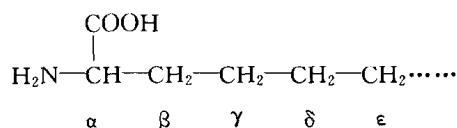
自从 1839 年德国化学家 G. J. Mulder 研究了乳和蛋中的清蛋白开始,至今已充分显示蛋白质是一切生命活动的主要物质基础。从某种意义上讲,蛋白质是生物大分子中最基本的也是最重要的,核酸要行使复制、转录或翻译等功能都离不开蛋白质;聚糖的功能也和蛋白质密切相关。所以,被命名为 proteins(源于希腊文中的 proteios, 意为 primary)。

蛋白质是细胞组分中含量最丰富、功能最多的生物高分子。由于其结构上的千变万化,功能也多种多样。机体内几乎无处不在,无处不发挥各种蛋白质特有的功能,如酶、抗原与抗体、多肽激素与其受体、各种转运蛋白、收缩蛋白以及细胞骨架和结构蛋白等等。由于蛋白质功能上的多样性是由其各自的独特结构所决定的,所以,只有在深入了解蛋白质结构的基础上才能更透彻地了解蛋白质的功能及其分子作用机制。

一、蛋白质的组成和一级结构

(一) 蛋白质的基本组成单位——氨基酸

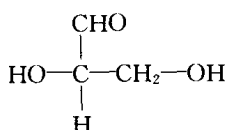
组成蛋白质的氨基酸除脯氨酸外,其余氨基酸均属 α -氨基酸(α -amino acid):



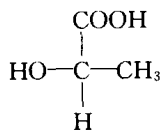
1. 氨基酸的构型

组成蛋白质的各种 α -氨基酸中,除了甘氨酸外,其 α 碳原子均属不对称碳原子,因而都具有光学活性。分子的手性(chirality)作为光学活性,以往一般用 D 或 L 代表构型。就各种 α -氨基酸而言,其连接在 C_α 四面角上各基团的排列与 L-甘油醛或 L-乳酸的构型相比较而得出其相对构型,均属于 L-氨基酸。

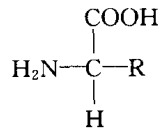
自然界中虽然存在一些 D-氨基酸(如细菌、细胞壁和某些抗生素),但在蛋白质组成中为什么均为 L-氨基酸?以 L-氨基酸参与合成蛋白质并不比 D-氨基酸有任何优越性,这或许与氨基酸发生过程中的机遇有关。然而,正因为 L-氨基酸参与组成了蛋白质,各种蛋白质的构象出现了一些现有的构型特征。



L-甘油醛



L-乳酸



L- α -氨基酸

自从 20 世纪 50 年代, Cahn, Ingold 和 Prelog 根据分子的手性特征提出了 RS 绝对构型的决定原则后, 氨基酸的构型也开始用 R 型或 S 型来表示(图 1-1)。这是根据不对称碳原子上连接基团的原子序大小而依次排列, 从远离最小基集团轴上观察其他基团, 原子序由大到小是按顺时针方向的属右手旋转, 称 R 构型 (rectus, 即右旋); 按逆时针方向的属左手旋转, 称 S 构型 (sinister, 即左旋)。为此, 由于基团原子序排列中 $\text{SH} > \text{OR} > \text{OH} > \text{NHR} > \text{NH}_2 > \text{COOR} > \text{COOH} > \text{CH}_2\text{OH} > \text{C}_6\text{H}_5 > \text{CH}_3 > \text{H}$, 因此, 在各种 L 型氨基酸中, 除了半胱氨酸为 R 构型外, 其余均为 S 构型。此外, 异亮氨酸及苏氨酸中 C_β 碳原子亦为不对称碳原子, 其绝对构型在异亮氨酸为 S 型, 在苏氨酸则为 R 型。

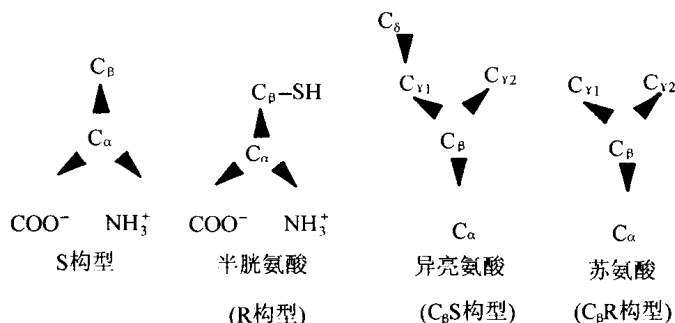


图 1-1 α -氨基酸的 R、S 构型

注: 按 Cahn-Ingold-Prelog 原则反向观察, H 在前

2. 生成蛋白质的氨基酸

根据生物体内蛋白质合成过程中的遗传密码, 可有 20 种 α -氨基酸参与, 这 20 种氨基酸称为编码氨基酸 (coding amino acid), 它们各有不同的侧链结构 R, 其性质有很大差异。即使同样含有的 α -氨基和 α -羧基, pK_a 值也有微小的差异。如组氨酸及苯丙氨酸的 α -羧基 pK_a 值为 1.8, 而苏氨酸可高达 2.6; 天冬酰胺的 α -氨基 pK_a 值为 8.8, 半胱氨酸的 α -氨基 pK_a 值可达 10.8, 显然是侧链结构的影响。于是, 通常将此 20 种氨基酸根据侧链的结构与性质的不同而分类(表 1-1):

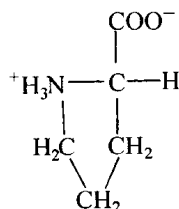
表 1-1 蛋白质中各种氨基酸基团的 pK_a 值

基团	蛋白质中 pK_a 值范围	游离氨基酸中 pK_a 值
α -羧基	3.5~4.0	1.8~2.6
侧链羧基	4.0~4.8	3.9(天冬), 4.2(谷)
咪唑基	6.5~7.4	6.0
巯基	8.5~9.0	8.3
酚羟基	9.5~10.5	10.1
α -氨基	8.0~9.0	8.8~10.8
侧链氨基	9.8~10.4	10.0
胍基	9~12	12.5

(1) 脂肪族氨基酸: 包括甘氨酸、丙氨酸、缬氨酸、亮氨酸及异亮氨酸。按照其

结构,氨基酸侧链趋于增大时,其疏水作用也趋于增强,即缬氨酸、亮氨酸及异亮氨酸远较甘氨酸、丙氨酸难溶于水。疏水作用强的氨基酸在蛋白质中常处于分子内或生物膜的疏水环境中。通常将脯氨酸也归入此类,实质上它属于亚氨基酸,其侧链已与其 α -氨基连接成环状。

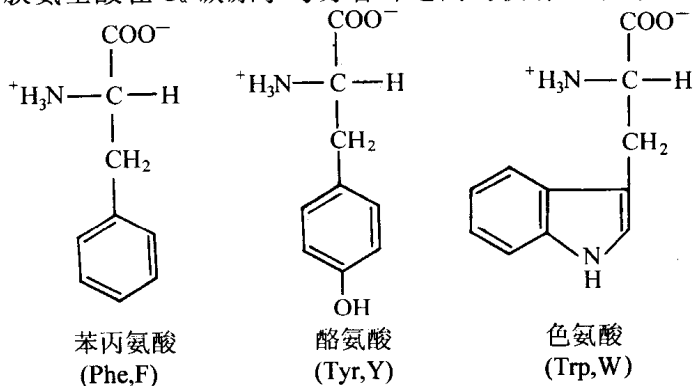
甘氨酸是惟一无侧链、也无不对称碳原子的氨基酸,在蛋白质的肽键平面之间不具有空间位阻,双面角可自由转动,以至肽链主链可任意弯曲,同时也能使蛋白质分子中肽链之间紧密靠拢。由于其柔性大,常出现在一些需要运动或转折的肽链片段之中。丙氨酸的侧链仅是甲基,虽属非极性,因其很小,在蛋白质分子中既可处于内部,也可处于表面,在蛋白质组成中一般较丰富。缬氨酸、亮氨酸及异亮氨酸以其侧链 R 上均带有分支而合称为支链氨基酸,其侧链均较大,有一定的空间位阻而限制了肽链的柔性,例如缬氨酸在 C_{β} 原子上分支,其 C_{γ} 甲基对主链有位阻,异亮氨酸也是如此;但亮氨酸的侧链分支出现在 C_{γ} 碳原子上,因此,对主链的空间位阻就小得多。不过,任何带有分支侧链的氨基酸都易于在蛋白质分子中固定于一定位置,并有利于链的折叠。脯氨酸由于氨基与侧链结合成环,具有固定的构型,以至 C_{α} 碳原子与肽键 N 之间的双面角只能固定在 $\pm 20^{\circ}$ 的范围之内,常常使肽链形成一个转角而改变主链方向,同时由于环的存在而具有空间障碍,对蛋白质二级结构如 α -螺旋及 β -折叠的形成都有破坏作用;但仍能出现在这些结构的末端或弯曲部位,因而常暴露于蛋白质分子表面,这与其他疏水性氨基酸有所不同。此外,此氨基酸还常以顺式肽键形式存在,为其特征。



脯氨酸(Pro, P)

(2) 芳香族氨基酸:芳香族氨基酸包括苯丙氨酸、酪氨酸和色氨酸。苯丙氨酸的疏水作用很强,而酪氨酸及色氨酸则由于侧链中带有极性基团,疏水作用较差,在高 pH 环境中,酪氨酸的酚基甚至还能解离出 H^+ 。

三种芳香族氨基酸在 C_{α} 碳原子与芳香环之间均仅有一个可转动的 C_{β} 甲烯基,



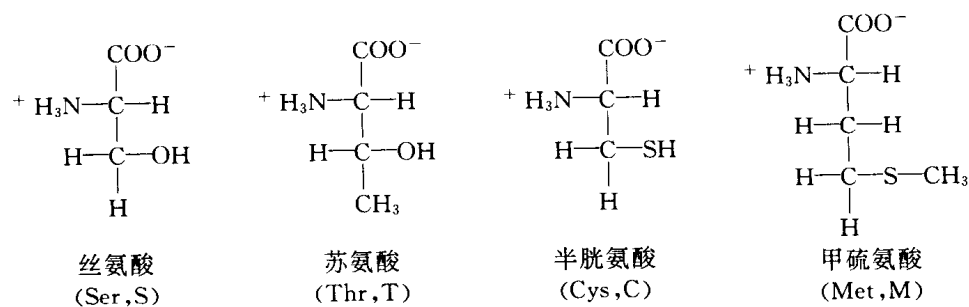
从而限制了侧链的柔性,尤其是色氨酸有较大的吲哚环更是如此。

酪氨酸酚基也可参与形成氢键,且较强。此类极化中性氨基酸虽有一定的亲水作用,但由于易形成氢键,故也常出现在分子内部。

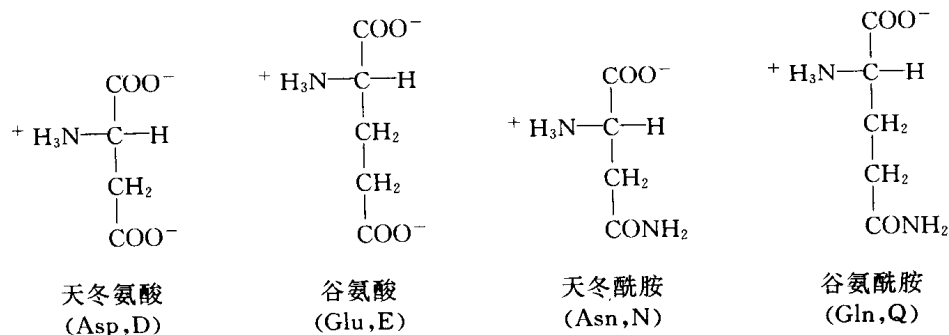
在蛋白质中,芳香族侧链环的堆积往往并不规则,实际上常以相互垂直的人字形排列,其存在形式对周围结构有很大影响。

(3) 羟基氨基酸与含硫氨基酸:羟基氨基酸与含硫氨基酸包括丝氨酸、苏氨酸、半胱氨酸及甲硫氨酸。由于侧链中带有极性基团,比相应的脂肪族氨基酸稍具亲水性。但甲硫氨酸疏水性仍很大,由于硫的参入,其柔性则较大。带有巯基的半胱氨酸,巯基并不易解离,但易氧化而形成带有二硫键的胱氨酸,这在蛋白质构象形成中具有重要意义。氨基酸残基带有羟基侧链者化学上比较活泼,常作为活性残基而具有十分重要的功能。同时,此种—OH基易与相邻的肽键基团—NH—或—C=O形成氢键。丝氨酸侧链短而苏氨酸侧链较长,有时可通过水分子而间接形成氢键。

半胱氨酸的巯基能与一些金属离子如Fe、Zn、Cu等配位结合,也能与结合蛋白质中一些非蛋白基团结合,甚至在蛋白质的空间结构中成簇地聚集,因此常埋藏在蛋白质分子内部,形成特定的基元[或称结构域单元、模体(motif)]结构。甲硫氨酸具有柔性疏水侧链,也常埋藏在蛋白质内部。



(4) 酸性氨基酸及其酰胺:天冬氨酸及谷氨酸属酸性氨基酸,无论游离状态或存在于蛋白质中,一般带负电荷,以酸根形式存在,在蛋白质内以盐键形式与正电荷基团连接。这两种氨基酸的酰胺,即天冬酰胺及谷氨酰胺,其侧链虽有极性,但不解离,也不带电荷。由于酰胺基中氨基易供氢,羰基易受氢,故这两种氨基酸在蛋白质中也易形成氢键。

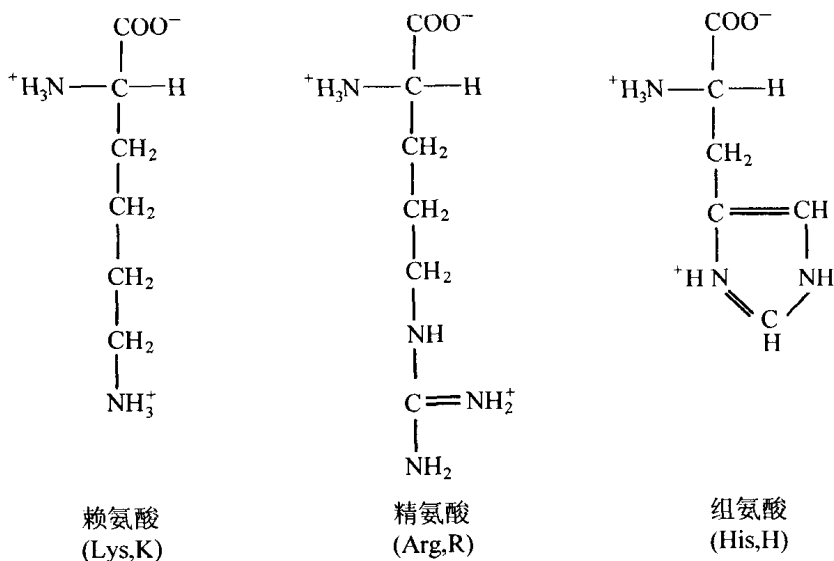


谷氨酸的带负电荷的 γ 羧基常位于蛋白质分子表面,能与钙结合,并且可在肽段序列中成簇存在,使分子表面形成不对称电荷分布,常常成为蛋白质与其他大分子物质相互作用的位点。

(5) 碱性氨基酸:碱性氨基酸包括组氨酸、赖氨酸及精氨酸,其中组氨酸的咪唑基 pK_a 值仅为 6.0,但在蛋白质中, pK_a 值可升高到 pH7 左右(表 1-1)。在生理 pH 条件下,常参与酶促反应中的质子传递反应,成为多种酶的活性中心。此外,也易与 Fe^{2+} 或其他金属离子形成配位化合物,参与组成各种金属蛋白质。

真正具有碱性的是赖氨酸及精氨酸,常结合 H^+ 而带正电荷,易与酸性基团形成盐键。它们一般存在于蛋白质分子表面,与带负电荷的酸性氨基酸残基形成盐键,也可与肽链 C 端形成盐键,甚至与带负电荷的核酸中磷酸基团形成盐键。因此,这两种氨基酸常涉及到蛋白质的活性。其中赖氨酸常与蛋白质的结合功能有关,精氨酸则与酶的催化功能有关,因为赖氨酸的长侧链可自由伸展在外,而精氨酸则可沿着侧链卷曲成疏水表面埋藏在蛋白质分子内,又可作为氢键供体易与蛋白质分子内的氧形成氢键而得以稳固所致。

总之,酸性及碱性氨基酸不仅带电荷,且极性较强。在蛋白质分子中,常出现在分子表面,与环境中的水分子结合而使蛋白质具有亲水性质,尤其是碱性氨基酸的侧链较长,柔性较大,在介质溶液中摆动,可增强其溶解度,同时也易作为多种酶的催化对象而出现在多种底物中。



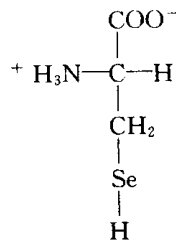
3. 蛋白质中的修饰氨基酸

蛋白质长肽链合成后或蛋白质翻译过程中,有些氨基酸残基经过修饰,改变其侧链的化学结构而出现的氨基酸叫做修饰氨基酸(modified amino acid)。常见的有 O-磷酸丝氨酸(或 O-磷酸苏氨酸、O-磷酸酪氨酸)、4-羟基脯氨酸(或 3-羟基脯氨酸)、 δ -羟基赖氨酸、 γ -羧基谷氨酸、甲基组氨酸(或甲基赖氨酸)、一碘酪氨酸(或二

碘酪氨酸、三碘甲腺原氨酸及甲状腺素)等,从而使原有的蛋白质出现了新的结构与功能,如 O-磷酸化氨基酸残基出现在一些蛋白质中,可以改变该蛋白质的生物学功能或酶活性。胶原中出现的羟基氨基酸残基是形成坚韧的胶原纤维的必要条件。 γ -羧基谷氨酸残基出现在蛋白质中,对于该蛋白质结合 Ca^{2+} 或通过 Ca^{2+} 与其他组分结合是十分重要的。甲状腺球蛋白中带有甲状腺素等碘化酪氨酸残基,显然是适应了甲状腺素合成及储存的需要。

除此之外,有些氨基酸在多肽链的 C 末端可经酰胺化修饰形成相应的酰胺而取消了羧基末端, N 末端经乙酰化形成一些乙酰基氨基酸残基而取消了氨基末端,谷氨酸在 N 端也可脱水形成焦谷氨酸残基而取消了末端氨基,这些修饰会影响蛋白质或多肽的功能或代谢。例如,氨基端经过上述修饰后,除了可保护蛋白质不受氨基肽酶作用外,还有利于该 N 端进入分子内外的非极性环境,完成一定的分子构象或参与功能活动。羧基末端修饰虽较少见(见于一些激素及蜂毒),其生物学意义可能与氨基末端修饰类似。

近年来,在含硒蛋白质中出现的硒代半胱氨酸(selenocysteine)并非翻译后修饰产物,而是在蛋白质翻译合成过程中,识别某些终止密码子 UGA 的特殊丝氨酸 tRNA,与丝氨酸结合后,可经酶催化丝氨酸而形成硒代半胱氨酸,再进入核糖体参与某些蛋白质的合成。由此可见,硒代半胱氨酸实为丝氨酸在蛋白质翻译合成过程中的修饰氨基酸,此种氨基酸在一些含硒蛋白质中具有重要的功能。



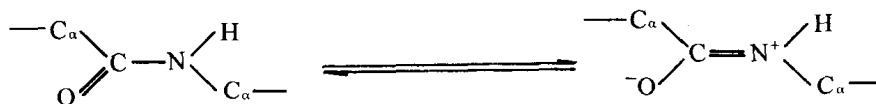
硒代半胱氨酸

(二) 氨基酸在蛋白质分子中的连接方式

在蛋白质合成过程中,主要的转变是氨基酸的 α 羧基与另一氨基酸的 α 氨基之间结合去水而形成酰胺键,通常称为肽键(peptide bond)。这是蛋白质主链的基本结构,所有的蛋白质实质上均属多肽链结构。肽链两端分别为游离的氨基端及游离的羧基端,在生理条件下分别带有正电荷及负电荷,加上肽链中可离解的侧链基团,各种蛋白质均属于多价两性电解质(polyampholytes)。随着蛋白质中酸性和碱性基团数及其 pK_a 值的差别,各蛋白质有其特征性的等电点。处于等电点时,蛋白质溶解度最小。不同蛋白质其等电点也不相同,在一定的 pH 环境中,所带电荷也有差异,这成为电泳、离子交换层析等分离蛋白质方法的重要依据之一。

在蛋白质或多肽的肽键中, $-\text{C}=\text{O}$ 与 $-\text{N}-\text{H}$ 之间呈反式而相互平行,很少扭曲。根据 1925 年 Linus Pauling 等对肽结晶中肽键各原子键长与键角的分析,发现形成肽键的 C 与 N 之间的键长比正常单键(0.149nm)短 10%,比 $\text{C}=\text{N}$ 正常双

键(0.127nm)又长,而C=O键长又较一般酮或醛中长0.002nm,认为C—N与C=O结构之间存在着共振,可测出的键长与键角不过是I型与II型两种形式结构的共振杂合体。因此,C与N之间具有半双键性质,从而形成酰胺平面。



I型结构中,C—N为单键,仅含轴对称的 σ 电子,故可自由旋转,不可能形成稳固的酰胺平面。II型结构中C=N为双键,含 σ 及 π 电子,因而有很大的偶极矩,不能旋转。而共振杂合体中约60%为I型,40%为II型, π 电子散布在C—O及C—N之上,形成半双键,实际测得的键长与键角如图1-2所示。

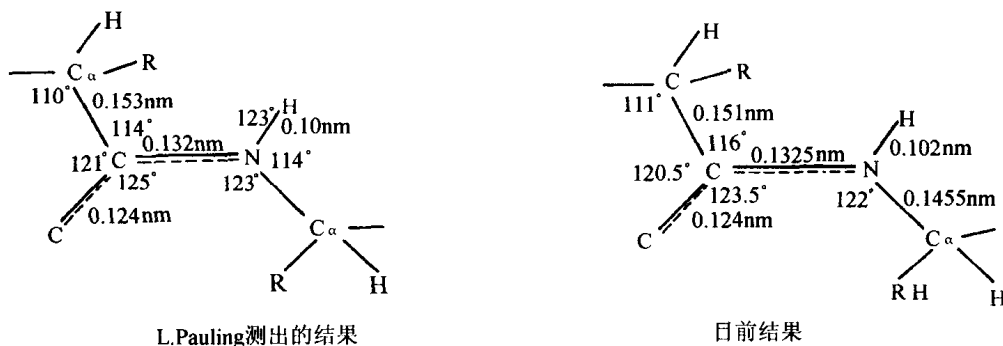


图 1-2 通常反式肽键中的键长与键角

虽然肽键也可能存在顺式构象,但由于靠近的侧链R往往会干扰顺式构象,故一般呈反式,仅在氨基酸羧基与脯氨酸的亚氨基结合成肽键时可出现顺式构象,这是由于脯氨酸残基吡咯烷环的 δ 碳原子在反式时反而能与邻近的氨基酸侧链发生空间排斥所致(图1-3)。

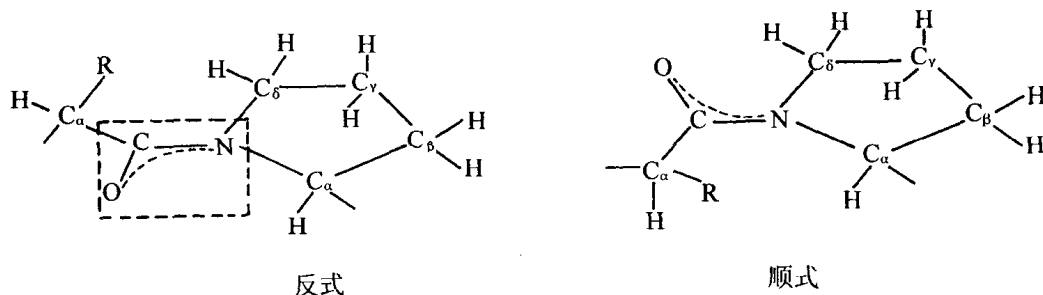


图 1-3 脯氨酸残基对酰胺平面构象的影响

(三) 蛋白质的一级结构

蛋白质的一级结构(primary structure)是指多肽链中氨基酸残基的排列顺序及二硫键所在的位置。通常由左至右,从氨基端开始,依次用氨基酸缩写符号向羧

基端书写。

1953年,英国 Sanger 花了大量的时间与人力,首先测定了胰岛素——一个最小蛋白质的一级结构,其相对分子质量是5 733,由 51 个氨基酸残基组成,含 A、B 两条链,A、B 链的氨基酸残基数分别为 21 及 30。

蛋白质的一级结构是其生物学活性及特异空间结构的基础。蛋白质之间的差别是由其氨基酸组成、氨基酸数目以及氨基酸在蛋白质多肽链中的排列顺序决定的。

蛋白质组(proteinome)是能体现正常细胞功能的整套蛋白质。

1. 种属差异

在蛋白质一级结构测定的基础上进行比较,对研究分子水平上蛋白质的生物进化产生了很有价值的成果。由于蛋白质的一级结构归根到底是基因遗传信息的表达,在进化过程中,随着时间的推移又以一定的概率进行着基因突变。这些突变往往可以在蛋白质一级结构中氨基酸残基的变化上反映出来,这在各种同源蛋白质一级结构的比较中清楚地显示出来。不同种属的生物中具有同一功能的蛋白质,在进化过程中可来自相同的祖先,但存在着种属差异。这是通过基因在进化过程中趋异突变(divergent mutation),导致蛋白质一级结构改变。如果影响其功能,则使生物体丧失其对环境的适应能力,就会被淘汰。有的突变不仅不影响该蛋白质的功能,反而导致生物个体的某种优势,形成新品种。此外,也有不影响生物功能的中性突变(neutral mutation),使该蛋白质仍保留下来,形成多种同源蛋白质的种属差异。这种差异在各种蛋白质分子的表面的氨基酸残基往往比较多见,而蛋白质分子内部残基的变异不仅少见,还往往借其他改变来进行代偿,以此来减少肽链主体的任何紊乱,而且变异也常常仅是类似性质的氨基酸残基,如赖氨酸转变为精氨酸,或异亮氨酸转变为亮氨酸等非极性氨基酸。此种变异并不一定与生物体形态或行为的进化相平行。例如,人的各种多肽链大约 99% 以上与黑猩猩的相应肽链是相同的,但形态及行为上却差别很大,这或许是由于生物进化不仅涉及结构基因突变而影响蛋白质结构,似乎调节基因的突变更为重要,而通过调节基因的变异仅是影响着各种基因的表达强度而已,并不涉及该蛋白质的一级结构。

细胞色素 C 是研究蛋白质一级结构种属差异最好的实例,因为至少有 120 多个生物种属的细胞色素 C 的一级结构已被测定。比较其结构可以发现,脊椎动物的细胞色素 C 均由 104 个氨基酸残基组成,昆虫及植物的细胞色素 C 常在氨基端再延伸一段。在各种细胞色素 C 中,几乎只有 28 个位置上的氨基酸残基是完全不变的。这些位置的氨基酸是维持其构象中发挥特有功能所必要的部位,属于保守性氨基酸残基。除此之外,其他各种非保守性氨基酸都可能随着进化而变异,而且在不同种属的细胞色素 C 的氨基酸差异数与种属之间的亲缘关系中,亲缘关系相近者,氨基酸差异少,反之则多。如马与酵母有 48 个残基不同;鸭和鸡有两个不同;人与恒河猴有 1 个不同;而人与黑猩猩完全相同。细胞色素 C 中氨基酸变异情况与生物进化的分类基本平行,故可据此描绘出生物进化的系统树。

2. 同源蛋白质

在蛋白质进化中更重要的问题是生物体中的同源蛋白质(homologous protein)。它们可以来自同一祖先蛋白质,随着进化而分化为具有不同功能特征的蛋白质。近年来,由于大量蛋白质一级结构的阐明,为此种蛋白质分化的研究提供了不少进展。例如,谷胱甘肽还原酶与硫辛酰胺还原酶均有转移氢到含硫化合物的催化功能,其一级结构中有40%以上位置的氨基酸残基相同。在胰凝乳蛋白酶原与胰蛋白酶原之间,或鸟氨酸转氨甲酰基酶和天冬氨酸转氨甲酰基酶之间,也有类似的情况。这是酶蛋白中某些氨基酸残基由于基因突变而变异,导致其底物专一性改变所致。此类变异无疑可以改变生物体的代谢方式,因此是蛋白质进化的重要问题。但是,在亲缘关系较远的蛋白质之间,由于相同残基较少,就难以决定是否属于同源蛋白质。事实上,任何两个无关的蛋白质肽链之间,一般都可能出现5%位置上残基相同的情况。在追踪蛋白质分子进化的谱系时,一般认为,至少必须有15%(去除间隔与插入)位置上残基相同的蛋白质才有可能属于同一家族的,15%~25%的残基相同者,属于准同源,必须进一步追索该蛋白质家族的“根”才行。近年来已发现了很多的同源蛋白质,特别是利用电子计算机储存了各种蛋白质一级结构的顺序以后,曾经发现血浆抗凝血酶Ⅲ不仅与 α_1 -抗胰蛋白酶的氨基酸顺序中有30%相同,两者又均为蛋白酶抑制剂,完全可能来自同一祖先而属于同源蛋白质。然而,蛋清中功能不明的卵白蛋白的一级结构也与以上两种蛋白质有30%同源,其间的关系就难以阐明了。

除了同源蛋白质趋异进化而导致进化,产生新功能的蛋白质而外,尚需考虑进化过程中同类蛋白质(analogous protein)的趋同变异(covergent mutation),导致产生类似的结构特征。这在目前,特别在亲缘关系较远的生物中也难以与同源蛋白质的趋异突变区分开来,两种进化显然是各有其不同的进化历程的。

Rossmann 根据一些酶 X 射线衍射图谱发现,许多结构上完全不同的蛋白质都含有一个约 70 个氨基酸残基形成类似构象的结构域,能结合单核苷酸辅酶,从而认为此种结构域可能是细胞前期原始蛋白质发生的雏形,似乎原始的且有功能的蛋白质是很小的,仅能与一些其他分子结合而已。当两个不同的小蛋白质结合起来,就能启动某种催化功能,再通过基因复制,就能发展成一个稳定的蛋白质家族,显然,蛋白质的进化是通过结构域基因的融合与增殖而进行的,这是比简单的基因突变更复杂得多的进化过程。

3. 蛋白质家族与超家族

通过快速而简便的测序方法,目前已积累了大量的蛋白质序列资料。根据数据库中收集到的蛋白质序列,至今已达 20 万左右,其中有的功能不明。就立体结构而言,仅有 8 000 种蛋白质已被测定,这对蛋白质分类产生很大困难,有必要从进化的角度上探求其来源而予以系统化处理。

1976 年, Dayhoff 根据蛋白质序列组成探求其同源关系,以研究其相关基因,阐