

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

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

出版

北京东黄城根北街16号

邮政编码:100717

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 插页:0

印数:1-5 000 字数:466 000

定价:34.00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 刘兴汉 初彦辉 李 晖

张 涛 张雪峰 范业鹏 欧 芹

杨春波 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, 一般是聚合体。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。

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。

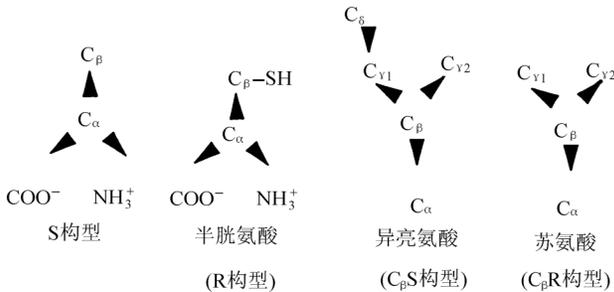


图 1-1 α -氨基酸的 R、S 构型

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):

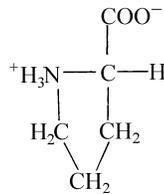
表 1-1 蛋白质中各种氨基酸基团的 pK_a 值

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。

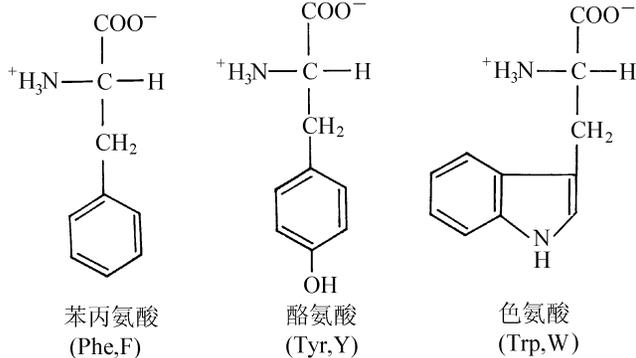
甘氨酸是惟一无侧链、也无不对称碳原子的氨基酸,在蛋白质的肽键平面之间不具有空间位阻,双面角可自由转动,以至肽链主链可任意弯曲,同时也能使蛋白质分子中肽链之间紧密靠拢。由于其柔性大,常出现在一些需要运动或转折的肽链片段之中。丙氨酸的侧链仅是甲基,虽属非极性,因其很小,在蛋白质分子中既可处于内部,也可处于表面,在蛋白质组成中一般较丰富。缬氨酸、亮氨酸及异亮氨酸以其侧链 R 上均带有分支而合称为支链氨基酸,其侧链均较大,有一定的空间位阻而限制了肽链的柔性,例如缬氨酸在 C_{β} 原子上分支,其 C_{γ} 甲基对主链有位阻,异亮氨酸也是如此;但亮氨酸的侧链分支出现在 C_{γ} 碳原子上,因此,对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 $\pm 20^{\circ}\text{C}$ 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} 亚甲基,



苯丙氨酸
(Phe, F)

酪氨酸
(Tyr, Y)

色氨酸
(Trp, W)

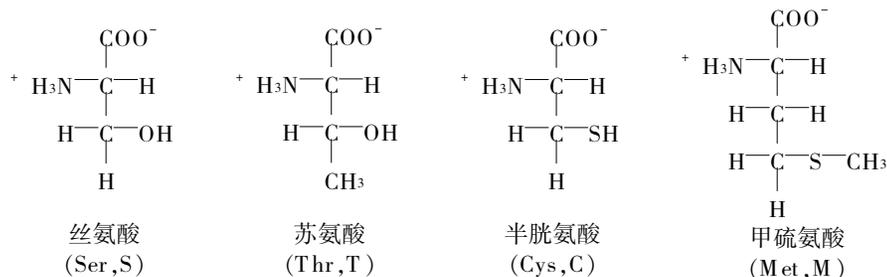
从而限制了侧链的柔性,尤其是色氨酸有较大的吲哚环更是如此。

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。

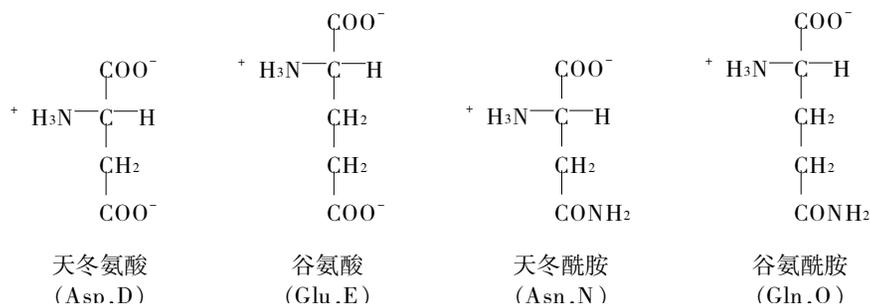
在蛋白质中,芳香族侧链环的堆积往往并不规则,实际上常以相互垂直的人字形排列,其存在形式对周围结构有很大影响。

(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。

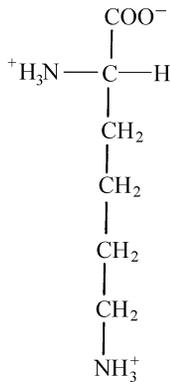


谷氨酸的带负电荷的 γ 羧基常位于蛋白质分子表面,能与钙结合,并且可在肽段序列中成簇存在,使分子表面形成不对称电荷分布,常常成为蛋白质与其他大分子物质相互作用的位点。

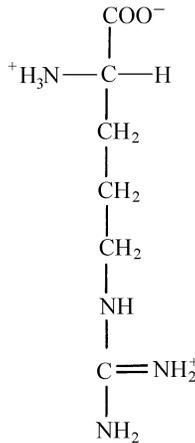
(5) 碱性氨基酸:碱性氨基酸包括组氨酸、赖氨酸及精氨酸,其中组氨酸的咪唑基 pK_a 值仅为 6.0,但在蛋白质中, pK_a 值可升高到 pH7 左右(表 1-1)。在生理 pH 条件下,常参与酶促反应中的质子传递反应,成为多种酶的活性中心。此外,也易与 Fe^{2+} 或其他金属离子形成配位化合物,参与组成各种金属蛋白质。

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。

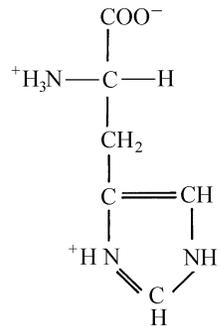
总之,酸性及碱性氨基酸不仅带电荷,且极性较强。在蛋白质分子中,常出现在分子表面,与环境中的水分子结合而使蛋白质具有亲水性质,尤其是碱性氨基酸的侧链较长,柔性较大,在介质溶液中摆动,可增强其溶解度,同时也易作为多种酶的催化对象而出现在多种底物中。



赖氨酸
(Lys,K)



精氨酸
(Arg,R)



组氨酸
(His,H)

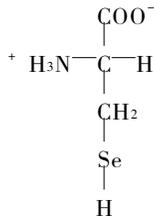
3. 蛋白质中的修饰氨基酸

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,可经酶促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。



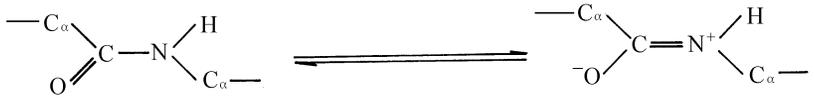
硒代半胱氨酸

(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,—C=O 与—N—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 C=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。

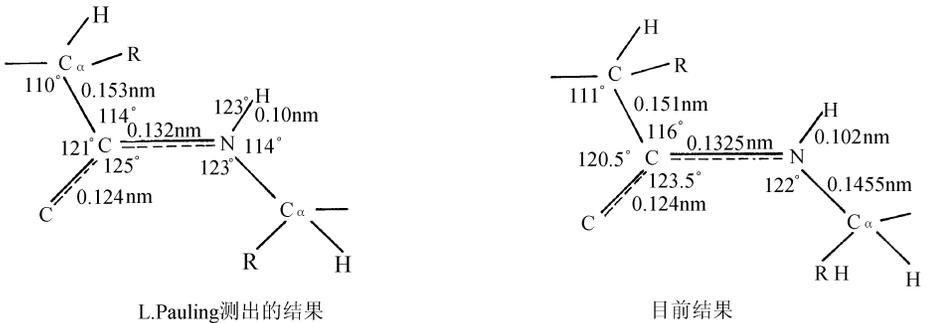


图 1-2 通常反式肽键中的键长与键角

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)。

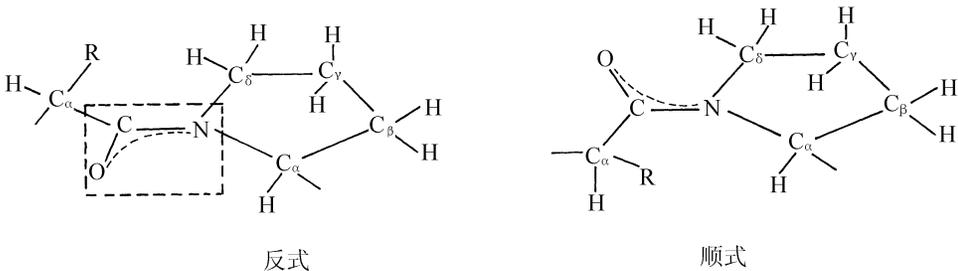


图 1-3 脯氨酸残基对酰胺平面构象的影响

(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常见,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 -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, Davhoff 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

明其进化历史。其序列中氨基酸残基组成少于 50% 的同源蛋白质,称之为超家族蛋白质,因其仍属同源,或者是同一祖先的复制产物,或者是其功能近似者,故名为超家族(superfamily)。至于不同生物中同源蛋白质因其氨基酸被替代、缺失,甚至是加入了一个或几个残基而产生的变异,其序列组成中残基的变化少于 50% 者,则属于同一家族蛋白质(protein family)。不同家族的蛋白质仍有相似的氨基酸序列,则组成超家族。因此,超家族中可能包含了几个蛋白质家族,超家族可视为为家族的联合,两者具有等级上的差别。如细胞色素 C 超家族中包含了细胞色素 C、C₆、C₂、C₅ 等 12 个家族。

然而,蛋白质通过不同结构域联合的进化观点必须对上述超家族概念进行一些修正。因为多结构域蛋白质可以有不同的结构域的不同进化来源历史,可以属于不同家族。如免疫球蛋白超家族中除了含免疫球蛋白相关结构域外,属于此超家族的 T 细胞受体中还结合了蛋白激酶结构域。还有许多参与动物细胞间粘接或细胞通讯的蛋白质的膜结合蛋白也含有免疫球蛋白结构域,往往在这些膜受体的细胞外部分为免疫球蛋白结构域,其细胞内部分则常常是激酶或磷酸酶,这就为超家族分类带来不少困难,必须从进化上分别探求其祖先,才能深入理解其进化来源与功能特征,单纯依据氨基酸序列异同的百分数就太片面了。已经发现细菌的伴侣蛋白(chaperonin)也有类似于免疫球蛋白的结构域,其氨基酸序列并不相似,然而从三维结构上证明其可能属于同一祖先。近年来已更多地考虑蛋白质立体结构中结构域或其结构域单元的相似性及其与功能进化的关系,来建立更为符合实际的蛋白质超家族概念。

在蛋白质家族中,一般根据其氨基酸序列相差少于 20% 者可再划分为多个亚家族(subfamily),参与 Ca²⁺ 结合的 EF 手蛋白家族现已分出了 31 个亚家族,如小白蛋白(或称副白蛋白)(parvalbumin)、肌钙蛋白 C、钙调蛋白、肠钙结合蛋白、肌浆钙结合蛋白等,由于其结构域来源有所区别,表现在各亚家族的功能及物理性质也有不同,但他们都有 EF 手结构来结合 Ca²⁺ 而完成不同的功能。

二、蛋白质的空间结构

各种各样的蛋白质几乎涉及所有的生物功能。生物功能的复杂性则有赖于蛋白质空间结构的复杂性及多样化。根据 1969 年国际纯粹与应用化学联合会(IUPAC)决定,蛋白质的空间结构统称为构象(conformation),即各种天然蛋白质分子中的多肽链并非以完全伸展的线状形式存在,而是通过分子中若干单键的旋转而盘曲、折叠形成特定的空间三维结构;蛋白质多肽链的所有原子由于单键的旋转而稳定于一定的空间排布,这与化合物的构型如氨基酸或糖的 D 或 L 构型不同,并不涉及原子之间的共价键的变化。

蛋白质的空间结构可以划分为二级、三级及四级结构。近年来由于空间结构研究方法的发展,有了不少进步。然而,彻底搞清复杂的空间结构仍然存在不少困难,至今已获得完全空间结构的蛋白质,与获得蛋白质的氨基酸序列相比,仍然为数甚

少。但是,已经探测了不少的空间结构规律,这有助于人们对蛋白质功能的深入认识,也有助于人们对蛋白质在漫长的生物进化过程中的发生与发展有所了解。

(一) 蛋白质的二级结构

1. 蛋白质的二级结构

蛋白质的二级结构(secondary structure)是指蛋白质多肽主链中原子的局部空间排列,并不涉及侧链的构象及与其他肽段的关系。它是指蛋白质多肽主链中有规则的重复构象。根据主链中各个酰胺平面之间出现二面角(ϕ, ψ)的不同,可以有不同的主链构象,即多种二级结构。然而要得到稳定的构象,主要依赖于酰胺键中的酰胺质子与羰基氧原子间形成较多的氢键。据此,L. Pauling 与 R. Corey 根据分子模型的构象于 1951 年提出了 α -螺旋及 β -折叠两种多肽主链的构象。

(1) 酰胺平面:肽键中的四个原子和与之相邻的两个 α -碳原子位于同一刚性平面,此刚性平面是构成肽链的基本平面,称为肽键平面,又叫肽单元(peptide unit)(图 1-4)。

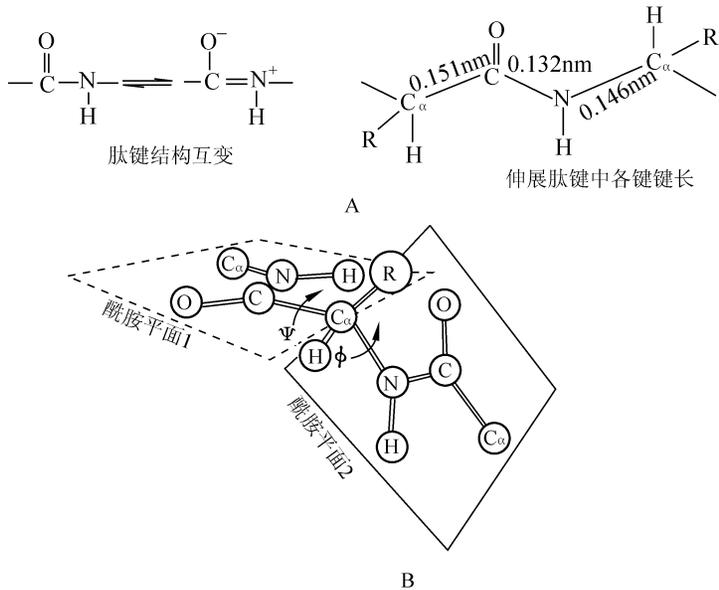


图 1-4 肽单元

(2) α -螺旋: α -螺旋(α -helix)的主链旋转方向,理论上既可是左手螺旋,也可以是右手螺旋。但由于组成蛋白质的氨基酸均为 L 构型,如果形成左手螺旋,则酰胺平面中的 C=O 将会与 L 构型的 C_{α} 所连接的侧链基团 C_{β} 太接近,以致发生空间干扰,影响其稳定性,故一般蛋白质中只出现右手 α -螺旋,在蛋白质的局部出现过少见的左手 α -螺旋(如组蛋白、羧肽酶 A 等);而有的蛋白质分子中含有右手螺旋很少,如肌动蛋白、 γ -球蛋白几乎不含右手螺旋。

在 α -螺旋中,多肽链主链骨架围绕中心轴螺旋式上升,每上升一圈为 3.6 个

氨基酸残基,相当于 0.54nm 垂直距离。因此,每个氨基酸残基沿中心轴上升 0.15nm , 旋转 100° , 相对酰胺平面的 ϕ 角为 -57° , Ψ 角则近于 -47° 。此时,上下螺旋圈之间形成较多的链内氢键,即一个酰胺平面上的 $\text{N}-\text{H}$ 中,氢原子可与其前面第三个酰胺平面的 $\text{C}=\text{O}$ 中氧原子形成氢键,而且氢键的方向与螺旋中心轴几乎平行。氢键所封闭的环由共价键及氢键构成,共包含了 13 个原子,设计 3.6 个氨基酸,故可用 3.6_3 螺旋来表示。在此种 α -螺旋情况下,每个氨基酸残基的高度恰巧相应于最适氢键的长度(约 0.28nm),同时酰胺平面中 N 、 C 及 O 与螺旋轴的距离分别为 0.157 、 0.161 及 0.176nm ,仅比通常的 van der Waals 半径大得很有限,有的还不到 0.01nm 。因此,整个 α -螺旋中,肽链主体的各个原子具有紧密堆积的结构特征,中心轴旁几乎不存在空间,因此,可以认为 α -螺旋中 3.6_3 螺旋是处于相当稳定的结构状态;而 2.27 、 3.0 、 4.416 等 α -螺旋稳定性较差。

然而, α -螺旋外侧 C_α 连接的 R 侧链仍然会影响 α -螺旋的稳定,甚至影响其形成。除了脯氨酸残基的吡咯烷环可以阻止 α -螺旋形成外,多肽链上连续存在的极性基团(如 Asp 、 Glu 、 Lys 等),它们之间的相互作用也严重影响螺旋的稳定性。即使是没有侧链的甘氨酸残基,也由于酰胺平面的自由旋转,二面角很难固定在 α -螺旋所需的范围内,也不易形成 α -螺旋。甘氨酸残基由于没有侧链,其 ϕ 及 Ψ 可以任意取值,因此,形成 α -螺旋需要的二面角的几率很小。

已知 α -角蛋白中大部分是 α -螺旋(α -helix)构成的长肽链骨架, α -螺旋也因此而得名。在脊椎动物中, α -角蛋白是毛发、皮肤及指甲、趾甲等结构中的主要蛋白质,其基本单位是角蛋白原纤维(protofibril)。目前认为,这是由两条双链 α -螺旋相互盘曲形成的左手超螺旋(superhelix),是一条四链结构。在每条单链中,虽然主要是右旋的 α -螺旋,但仍有少数非螺旋区(图 1-5),在超螺旋中,各链之间借助于紧靠

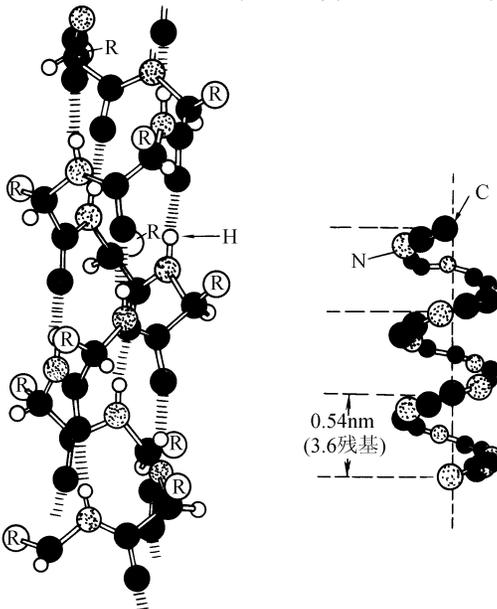


图 1-5 α -螺旋

的非极性侧链的 van der Waals 力而使其稳定。此外,肽链之间的二硫键也起了很强烈的作用。

(3) β 折叠:有些纤维状蛋白质的肽链处于比较伸展的状态,并不卷曲成螺旋,却以锯齿状片层结构存在,在折叠层中邻近肽链的骨架之间仍可通过 N—H 与 C=O 之间形成氢键,从而使多条肽链平行稳定排列,如扇面折叠状。此种二级结构称为 β 折叠(或片层)(β -pleated sheet)。通过此种锯齿状折叠,与 C_{α} 相连的侧链可以交替地位于折叠片层的上方和下方,从而避免了相邻侧链 R 之间的空间障碍,且可以形成最多的氢键,使 β 折叠得以稳定(图 1-6)。

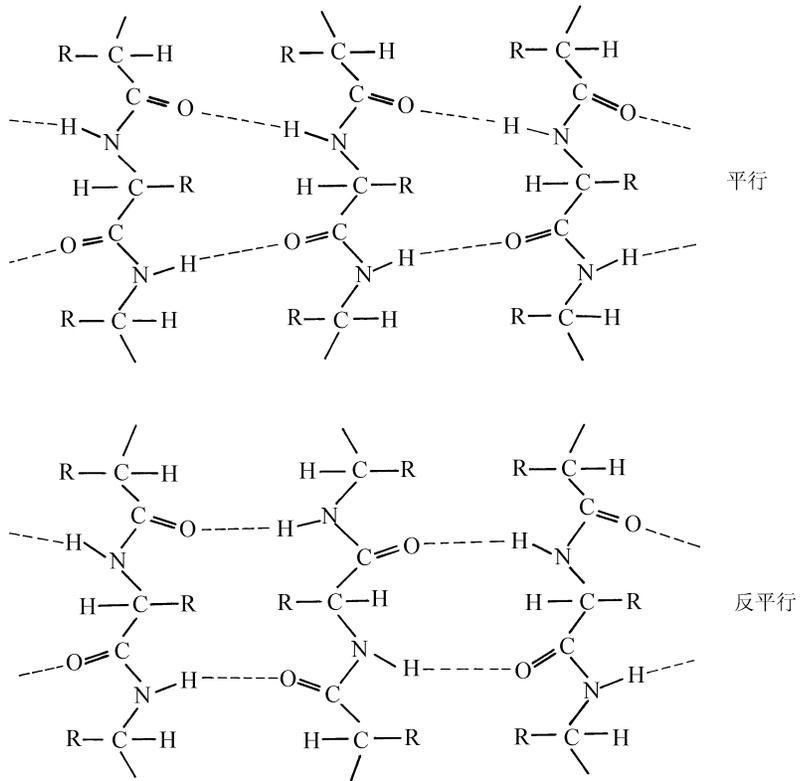
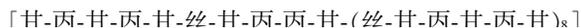


图 1-6 平行与反平行 β 折叠

上图:平行 β 折叠; 下图:反平行 β 折叠

β 折叠根据平行肽链主体之间走行方向可分为平行与反平行两种方式。如各个肽链 N 端均在 β 折叠的同一端则属平行式;如各个肽链 N 端按反方向排列,则为反平行式。平行式 β 折叠中酰胺平面的转角 ϕ 为 -119° , ψ 为 $+113^\circ$;而反平行式 β 折叠中的酰胺平面转角 ϕ 为 -139° , ψ 为 $+135^\circ$,以后者较为稳定。

丝心蛋白是蚕丝中主要蛋白质,其中大片的反平行式 β 折叠以平行方式堆积成多层结构。就其一级结构而言,很多区段为下列氨基酸序列的重复排列。



由此可见,整个氨基酸序列中,几乎都是隔位排列着甘氨酸残基,而甘氨酸残基之间则为结构简单的丙氨酸或丝氨酸残基,这暗示着甘氨酸残基将全部位于 β 折叠层的同侧,而丝氨酸及丙氨酸残基位于 β 折叠层的另一侧,从而使交替叠成 β 折叠层之间分别是甘氨酸残基聚集区和丙氨酸(及丝氨酸残基)聚集区,以致 β 折叠层之间的距离分别为0.35nm及0.57nm。此外,由于缺少半胱氨酸残基, β 折叠之间不是用二硫桥交联的,而在其间仅由van der Waals力连接着,以致丝绸可以显出很柔软的性质。但是, β 折叠中多肽链处于比较伸展的情况, β 片层中肽链之间又有很多氢键,虽不能再过度伸展,但能承受相当大的张力而显得相当牢固(图1-7)。

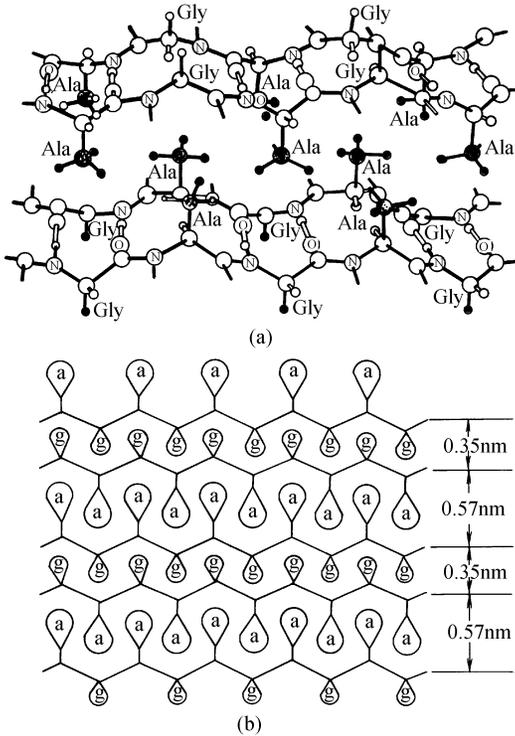


图 1-7 丝心蛋白中 β 片层间的残基排列
(a) 原子模型; (b) 片层堆积方式

除了以上三种氨基酸残基外,丝心蛋白中尚含有少量缬氨酸及酪氨酸等残基,并不参与 β 折叠的结构而在丝纤维中间隔存在。由此形成的不规则构象或许有利于丝纤维的伸展而稍具弹性,不同种类蚕丝伸展程度不同,可能正是少数这类氨基酸组成有所差别所致。

(4) β 转角:球状蛋白质的紧密球形源于肽链走向的多次逆转,主要是通过 β 转角来实现的。因此, β 转角(β -bend)是球状蛋白质的重要二级结构,在其他各种二级结构之间起连接作用。它可出现在 α 螺旋之间, α 螺旋与 β 折叠之间,或反平行 β 折叠的片层之间连接两段肽链。在球状蛋白质含量颇多,并常位于分子表面。

这与亲水性氨基酸残基形成 β 转角的倾向很强有关(表 1-2)。此外,甘氨酸及脯氨酸由于其结构特点也常出现在 β 转角中,在 β 转角的中间两个残基中几乎有 2/3 是脯-天冬-NH₂ 或脯-甘残基对。

表 1-2 氨基酸残基形成 α -螺旋、 β -折叠及 β 转角 倾向的概率(P)

氨基酸	P_{α}	α -螺旋形成倾向	P_{β}	β -折叠形成倾向	P_t
谷	1.53	+++	0.26	--	0.44
丙	1.45	+++	0.97	+	0.57
亮	1.34	+++	1.22	++	0.53
组 ⁺	1.24	++	0.71	-	0.69
甲硫	1.20	++	1.67	+++	0.67
谷-NH ₂	1.17	++	1.23	++	0.56
色	1.14	++	1.19	++	1.11
缬	1.14	++	1.65	+++	0.30
苯丙	1.12	++	1.28	++	0.71
赖 ⁺	1.70	+	0.74	-	1.01
异亮	1.00	+	1.60	+++	0.58
天冬 ⁻	0.98	0	0.80	0	1.26
苏	0.82	0	1.20	++	1.00
丝	0.79	0	0.72	-	1.56
精 ⁺	0.79	0	0.90	0	1.00
半胱	0.77	0	1.30	++	1.17
天冬-NH ₂	0.73	-	0.65	-	1.68
酪	0.61	-	1.29	++	1.25
脯	0.59	--	0.62	-	1.54
甘	0.53	--	0.81	0	1.68

注：“+”表示形成倾向；“-”表示破坏倾向；“0”表示差别不大

β 转角可使蛋白质分子中肽链出现 180° 回折,同时常在第一个残基的 $C=O$ 与第四个残基的 $N-H$ 间形成氢键,成为 4 个连续残基的稳定结构。 β 转角有 I 型 (β_I) 和 II 型 (β_{II}) 两种形式。在球状蛋白质中普遍存在的普通转角 (β_I), 其双面角 ϕ 为 -60° , Ψ_2 为 -30° , ϕ 为 $+90^\circ$, Ψ_3 为 0° , 4 个 C_α 不在同一平面上而有 $+45^\circ$ 双面角。 β_{II} 型转角又称甘氨酸转角, 其 ϕ 为 -60° , Ψ_2 为 $+120^\circ$, ϕ 及 Ψ_4 与 β 相同, 使 4 个 C_α 与氢键几乎在同一平面上, 此种转角中第二个残基大多为甘氨酸, 故名甘氨酸转角(图 1-8)。还有一些非典型的 β 转角, 可以有两个氢键, 或者没有氢键而靠邻近侧链间的作用, 乃至远程作用力来维系其转角结构。

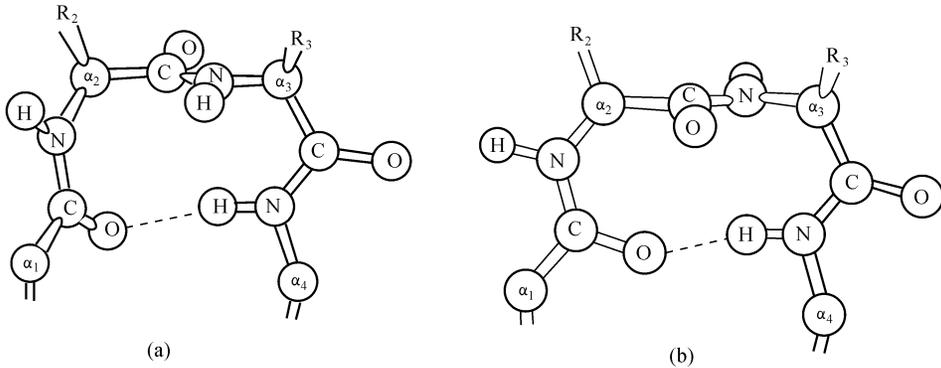


图 1-8 两种主要 β 转角结构示意图
(a) β_I (普通转角); (b) β_{II} (甘氨酸转角)

(5) 无规卷曲: 无规卷曲(random coil)原指没有确定规律性的肽链构象, 但仍然是紧密有序的稳定结构, 可通过主链间氢键, 甚至主链与侧链间形成氢键而维持其构象。球状蛋白质中含量很多, 类型大小不一。已发现的大体上分为两大类。

1) 紧密环(compact loop): 肽链片段的主体结构形成一端开放的环状, 其主链卷曲成希腊字母 Ω , 故又称 Ω 环(omega loop)。此肽链片段由 6~16 个氨基酸残基组成, 显然大于 β 转角, 其残基侧链可堆积在环内, 通过疏水作用或形成氢键等, 成为紧密的结构, 肽链片段两末端间距离小, 从第一个 C_α 到最后一个 C_α 之间小于 1nm, 常为 0.37~1.0nm。就 Ω 环的氨基酸残基组成而言, 与 β 转角相似, 甘、脯、天冬、天冬-NH₂ 及丝氨酸残基较常见, 也常出现酪氨酸残基。亲水性残基较多, 故此环常出现在球状蛋白质分子表面。有的蛋白质带有多个 Ω 环, 在空间结构中常成簇存在。大的蛋白质中有多个环, 有的可占残基总数的 21%。 Ω 环可能与蛋白质的识别功能有关, 并参与催化作用。从蛋白质数据库中已发现近千个 Ω 环。

2) 连接条带: 伸展的肽链条带(straps)构成二级结构单元之间的连接, 有以下几种类型: ① S 形的 α - α 连接; ② 长度及构象不一的 α - β 和 β - α 连接; ③ β 弓连接了不相邻的 β 链, 形成“希腊花边连接(Greek key connection)”(图 1-9)。

各种连接条带的长度、走向颇不规则, 可以反向改变肽链的走向, 也可改变为 90° 走向, 有些使肽链微微弯曲, 可使肽链密集, 也可发生扭结, 在蛋白质肽链的卷

曲、折叠过程中具有明确的结构作用。

连接条带主链 CO 和 NH 之间一般不形成氢键。对不同种属的同源蛋白质的残基序列进行比较,常有残基在连接区出现插入、替换、消除的变异,说明在蛋白质进化过程中,连接肽段具有易致突变的情况。从功能上看,此类连接条带往往参与形成蛋白质的一些结合位点或酶的活性位点。

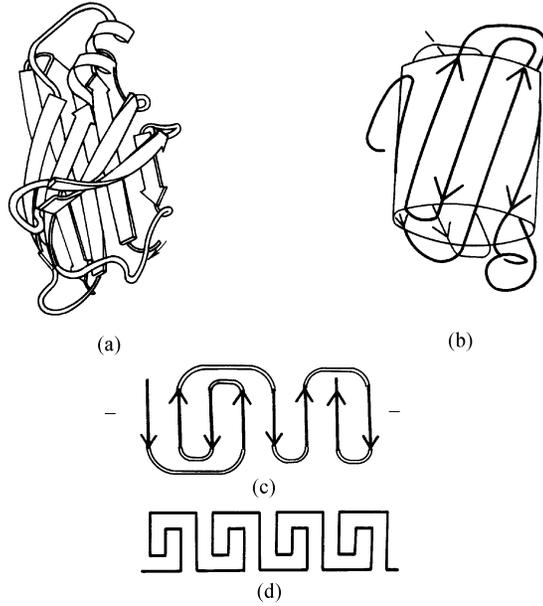


图 1-9 前白蛋白单体的希腊花边 β 桶结构

(a) 前白蛋白构象; (b) 前白蛋白三维拓扑图; (c) 摊开后的二维; (d) 希腊花边图案

(6) 无序结构:从 X 射线衍射分析发现,蛋白质分子中还确实存在着没有确定空间结构的区域,这种无序结构是由于其不断运动,或是该区域具有不同的构象,以致 X 射线衍射得不到结构图像。蛋白质分子中暴露在介质中带电荷长侧链的残基如赖、精、异亮氨酸等常是无序的,肽链末端的几个残基如果没有较大的疏水性氨基酸,也常表现为无序结构。

无序结构的生物学意义也不确定,有时切掉肽链末端的无序残基并不影响该蛋白质的活性,但在有些蛋白质的配体结合部位,无配体时呈无序结构,一旦结合上配体则成为有序结构。

2. 超二级结构

(1) 超二级结构:在多肽链内顺序上相互邻近的二级结构常常在空间折叠中靠近,彼此相互作用,形成规则的二级结构聚集体或二级结构簇构象,它是介于二级结构和结构域之间的一个构象层次,称为超二级结构(supersecondary structure)。它可直接作为三级结构“建筑块”或结构域的组成单位,是二、三级结构间的

一个层次。在球状蛋白质中,几乎所有的三级结构都可以用这些折叠类型,乃至它们的组合型来予以描述,因此,又将其称为标准折叠单位(standard folding units)。常见的标准折叠单位有 $\alpha\alpha$ 、 $\beta\beta$ 、 $\beta\alpha\beta$ 三种: $\alpha\alpha$ 是相邻的两条或三条 α -螺旋通过肽链连接而成左手超螺旋结构; $\beta\beta$ 是 β -折叠链逆转成发夹结构联结而成回形(greek key)拓扑结构; $\beta\alpha\beta$ 则由一个不规则的环链或一个 α -螺旋将两条平行的 β -折叠末端以右手或左手交叉方式(图 1-10)连接起来。它们所产生的构象形式既与所连接的二级结构单元种类有关,也与连接肽链的长度、残基构象或其极性、带电性及疏水性有关,可以出现两条 α -螺旋弯成 90° 角的 α -拐角(α - α corner),两条 α -螺旋平行的 α -发夹,连接不同类型的二级结构单元则可出现拱形结构等。

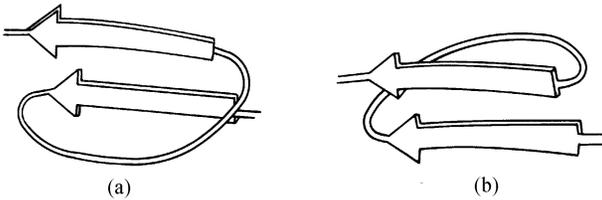


图 1-10 平行 β 链的右手及左手交叉连接

(a) 右手交叉连接; (b) 左手交叉连接

(2) 结构域:结构域(domain)是超二级结构与三级结构之间的一个层次。在较大的蛋白质分子中,由于多肽链上相邻的超二级结构紧密联系,形成两个或多个在空间上可以明显区别的局部区域。结构域与分子整体以共价键相连,一般难以分离,这是它与蛋白质亚基结构的区别。一般每个结构域约有 100 ~ 200 个氨基酸残基组成,各有独特的空间构象,并承担不同的生物学功能。如果短的多肽链仅有一个结构域,则该蛋白质的结构域和三级结构为同一结构层次。而较大的蛋白质常含有两个或两个以上的多个结构域。

3. 三级结构

具有二级结构、超二级结构或域结构的一条多肽链,由于其序列上相隔较远的氨基酸残基侧链的相互作用,而进行范围广泛的盘曲与折叠,形成包括主、侧链在内的空间排列。这种在一条多肽链中所有原子在三维空间的整体排布称为三级结构(tertiary structure)。蛋白质的三级结构是每一条多肽链内全部二级结构的总和及所有侧链原子的空间排布和它们的相互作用关系。可以从不同层次上来探索蛋白质的三级结构规律。

根据组成结构域的二级结构类型和组合,结构域大致分为 4 种主要类型,大多数蛋白质或其结构可归属这些类型中。

(1) 全 α 类型(或反平行 α):结构中主要的二级结构是 α -螺旋。各段 α -螺旋通过反平行或近于相互垂直的状态而连接,排列成层,再平行叠起,从而形成三级结构的构象基础。还可以进一步细分为几个亚型:

1) 升降螺旋捆(up and down helix bundle):由各邻近的 α -螺旋头尾相连,反向排列,形成筒状(图 1-9a)。常见的为四螺旋捆,其中 α -螺旋是不扭转也不弯曲而排列的,如细胞色素 b₅₆₂(图 1-11);而结合在 RNA 的引物阻抑物(repressor of primer, Rop)则在 α -螺旋之间扭成超螺旋。其他如载脂蛋白 E₃、人生长激素、去铁铁蛋白单体、细胞因子如白介素-2、白介素-4、粒细胞-巨噬细胞集落刺激因子,乃至导致神经退化变性的朊病毒(pinon),都属于四螺旋捆类型,也有蛋白质呈三螺旋捆或较大的五或六螺旋捆。

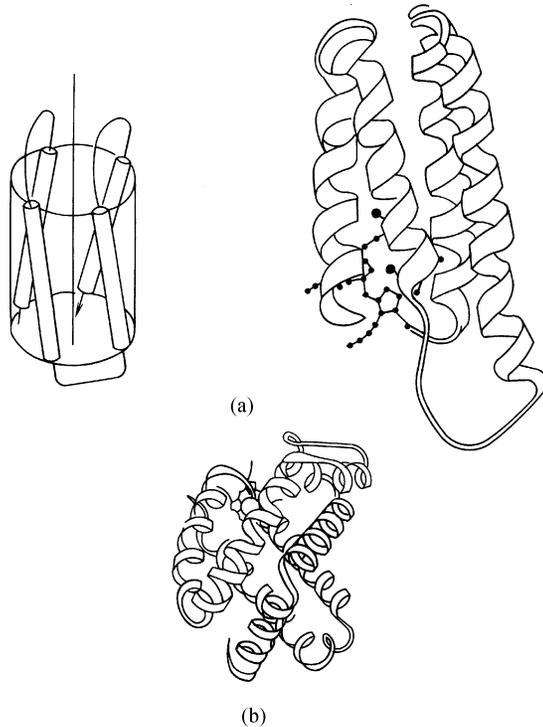


图 1-11 全 α 型三级结构类型

(a) 升降螺旋捆模式及细胞色素 b₅₆₂; (b) 血红蛋白 β 亚基的希腊花边螺旋捆结构

2) 希腊花边螺旋捆(Greek key helix bundle):多个 α -螺旋段近于垂直地相互连接起来,并卷曲成形似古典希腊花瓶上见到的“回”字形花边(图 1-9b~ d)。如肌红蛋白或血红蛋白亚基,一般常是两层 α -螺旋叠起的结构,但也有蛋白质是叠成三层的。

(2) 平行 $\alpha\beta$ 类型:整个肽链中 α -螺旋与 β -折叠交替存在,卷曲组成多层。一般平行的 β -折叠片层在结构域的内部, α -螺旋覆盖在外部,实际上是超二级结构 $\beta\alpha\beta$ 的进一步连接扩充,大多卷曲成右手交叉连接。如平行 β 桶(parallel β -barrel)结构,多见为 8 条平行 β -链在中央,相互连成氢键网,排列成内桶;与各 β -链连接的 8 条 α -螺旋则在其间沿着相同的方向排列在 β -桶的外侧,成为外桶;整个 β -桶结构为 $(\beta\alpha)_8$ 。此类结构颇为多见,大约 10% 的酶分子结构属于此类。桶高低则不一致,催

化的反应也多种多样,如磷酸丙糖异构酶、木糖异构酶、丙酮酸激酶等。有些有一定的氨基酸残基序列同源性,故有人相信他们是从同一老祖先进化而来(图 1-12a)。有些蛋白质的 β 桶结构为 $(\beta\alpha)_7$ 或 $(\beta\alpha)_{10}$ 不等。

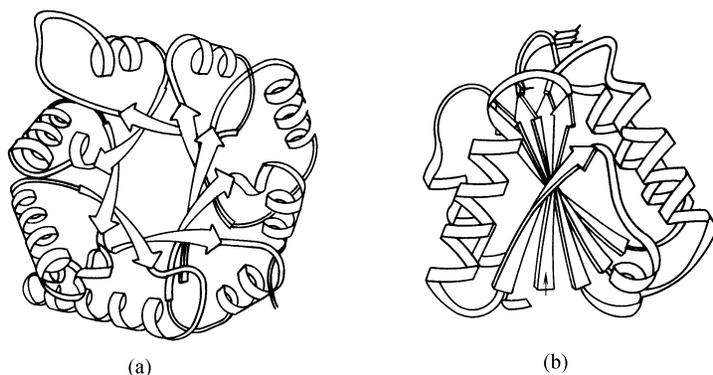


图 1-12 平行 $\alpha\beta$ 三级结构类型

(a) 磷酸甘油醛异构酶的 β 桶结构域(顶面观); (b) 黄素氧还蛋白的核苷酸结合结构域

常见的核苷酸结合结构域是由许多平行的 β -折叠片层形成马鞍形排列, β -链之间由 α -螺旋形成右手交叉连接,排在 β 片层两侧。而 β -片层末端则有核苷酸结合位点,可以结合 NAD^+ 、FMN 或 ATP 等(图 1-12b)。各种以 NAD^+ 为辅酶的脱氢酶,甚至以 NADP^+ 及 FAD 为辅酶的蛋白质中都有此结构域。此外,腺苷酸激酶、 F_1ATP 酶、肌球蛋白、延长因子 EF-Tu、癌基因 Na-ras 产物 P21^{ras}、己糖激酶等也有此结构域。

(3) 反平行 β 类型: 主要由 β -折叠片层反平行排列而形成。 β -链之间的连接以 β 转角,或以跳过相邻 β -链(1~4 个 β 链)的条带而进行连接。常见的拓扑结构希腊花边 β 桶(亦称回波纹)(图 1-9)的 β -链呈逆时针方向盘绕,桶的相对两边 β -链呈右手交叉连接,如超氧化物歧化酶、免疫球蛋白 V_L 结构域就是如此。少数蛋白质仅为升降 β 桶结构,如视黄醇结合蛋白、黄豆胰蛋白酶抑制剂等,形似果酱卷饼。

反平行 β -折叠片层也可排列成单层,不成桶状,称为开面夹心(open face sandwich)结构,其一侧可有一些 α -螺旋或环的组成,也有的一侧暴露于溶剂中,与其他结构域结合。细菌叶绿素蛋白具有典型的开面夹心结构(图 1-13c)。

(4) 模体: 模体(或称结构域单元、基元、基序)是结构域的亚结构,可体现结构域的各种生物学功能,但氨基酸序列较短。随着蛋白质三级结构研究的增多,发现不少蛋白质分子的空间结构中存在着某些立体形状或拓扑结构颇为类似的局部区域,可称之为模体(modif)。1973 年, S. T. Rao 与 M. G. Rossmann 就提出黄素氧还蛋白(flavodoxin)与乳酸脱氢酶结构中具有类似的 $\beta\alpha\beta\alpha\beta$ 超二级结构,其二级结构单元的安排及连接的拓扑结构非常近似,但连接环的大小及构象并不相同,此即以后在许多蛋白质中出现的核苷酸结合结构域,可称之为核苷酸结合模体(图 1-12b)。在许多并无氨基酸序列同源性的不同蛋白质中都可出现。不久有人发

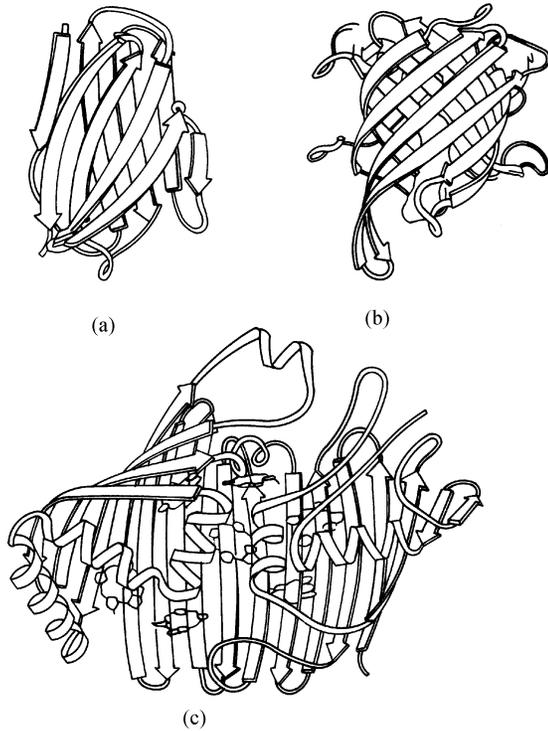


图 1-13 反平行 β 三级结构类型

(a) 免疫球蛋白 V_L 结构域; (b) 视黄醇结合蛋白; (c) 叶绿素蛋白的开面夹心结构

现,超氧化物歧化酶(SOD)与免疫球蛋白的 β 桶结构中也有拓扑结构相同的连接类型,即希腊花边 β 桶模体(图 1-9)。以后又陆续出现了平行 β 桶模体(图 1-12a)、升降螺旋捆模体(图 1-11a)等存在于多种不同的蛋白质结构中,其间也难以找出结构在进化上的同源关系。

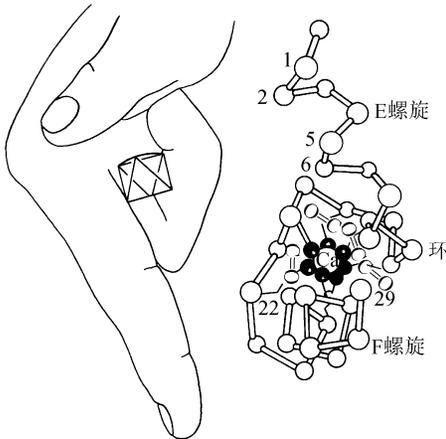


图 1-14 EF 手钙结合模体示意图

近年来,借助于磁共振等技术可以测得溶液中蛋白质的小结构,通过蛋白质的氨基酸序列信息在电子计算机上对比,从而得到许多小型模体的结构类型,通常由 20~ 30 个氨基酸残基组成,常有特殊的结合功能,可在不同的蛋白质中出现,甚至重复出现。如 40 种钙结合蛋白中存在着约由 12 个氨基酸残基的钙结合环,连接在两段 α -螺旋之间,形成螺旋-环-螺旋(HLH)钙结合模体,形状似拇指与示指(E 及 F 螺旋)直角相交,将 Ca^{2+} 配位结合(环),故又称 EF 手模体(图 1-14)。

不少蛋白质中存在卷曲螺旋型(coiled coil) α -螺旋模体,可以是两条、三条或四条右手 α -螺旋通过左手卷曲而形成的超螺旋,二链间约为 20° 夹角不等, α -螺旋间可以通过疏水性残基相互结合。因此,在重复的由七氨基酸残基形成的 α -螺旋中第1、4氨基酸呈疏水性,而使两条 α -螺旋相互借疏水性交叉连接,亮氨酸拉链(leucine zipper)模体(图 1-15)就是如此。真核细胞的许多转录因子含有此种类型模体,由两条平行 α -螺旋链组成。

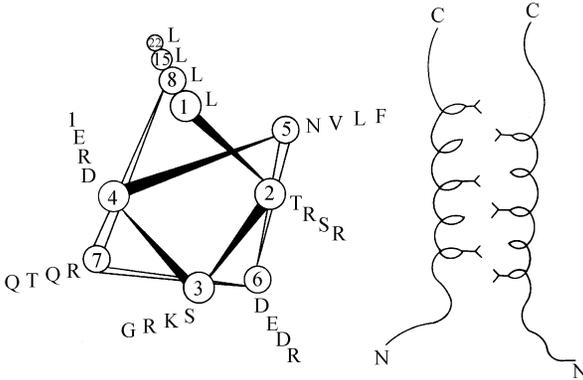


图 1-15 亮氨酸拉链模体示意图

有些 DNA 结合模体为反平行螺旋-转角-螺旋(HTH)的卷曲螺旋,其中第二条螺旋常在 DNA 双螺旋的大沟中与碱基特异结合,成为识别螺旋。虽然根据 α -螺旋结构、定位及转角的组成情况还可细分为多种,但模体结构形状及其与 DNA 结合方式还是十分相似的。在真核转录因子常出现 HLH 模体,是由 5~ 24 个残基组成的环连接在两个 α -螺旋之间,螺旋中疏水性残基位置也十分保守,易使蛋白质肽链之间结合形成二聚体疏水核心,再参与 DNA 结合。至于形成三链卷曲螺旋的则有层连蛋白、纤维蛋白原、血影蛋白、 α 辅肌动蛋白及甘露糖结合蛋白等。四条 α -螺旋卷曲成四链螺旋捆模体更为多见,常是几条多肽链缔合而形成的,在肿瘤抑制物 P₅₃、lac 阻抑物及细菌荧光素酶等中都有出现。载脂蛋白 E₃、人生长激素及人白介素-4 则有单条肽链形成的四螺旋捆模体。有的蛋白质通过亚基的聚合还可形成更大的螺旋捆,如大肠杆菌的延胡索酸酶 C 是由 4 个五螺旋捆聚合而成的 20 个 α -螺旋的大模体核心,大肠杆菌的天冬氨酸受体则是由两个亚基中的反平行四螺旋捆相互接合而形成的二聚体,此种结构在亚基聚合中显然发挥了重要的作用。

由两条反平行 β -折叠链通过 2~ 5 个残基连接而形成的 β -发夹模体是蛋白质中常见的构象单位。在与基因表达调控有关的 DNA 结合蛋白中常常重复出现的典型锌指模体(图 1-16),正是由于 β -发夹中两个半胱氨酸残基与连接在 β -发夹后的 α -螺旋上两个组氨酸残基共同与使结构稳定的锌原子配位结合,成为 $\beta\beta\alpha$ 折叠构象 DNA 结合功能。有些锌指模体则为 $\beta\beta\beta\alpha$ 或更为复杂的构象,或是通过四个半胱氨酸残基与锌原子配位结合而形成。除了一些转录因子外,在一些类固醇激素受体乃至艾滋病病毒核壳蛋白也存在锌指模体,有些还能与 RNA 结合以发挥其作用。

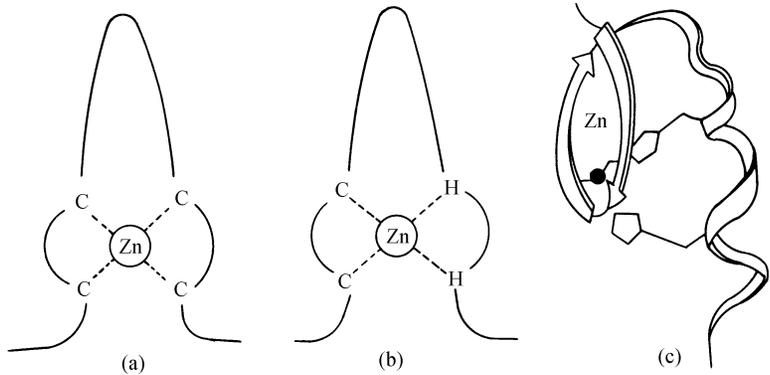


图 1-16 锌指模体示意图

(a)、(b)示两种锌指组成；(c)示 $\beta\alpha$ 类型的锌指模体

此外,反平行的 β -发夹模体还可进一步组合成其他类型的复杂模体,如希腊花边模体。奇特的 β -螺旋(β -helix)模体是由许多平行的 β 片层在肽链中依次卷曲形成宽大的螺旋圆柱构象。此种模体首先在细菌的果胶裂解酶 C (pectate lyase C) 中发现,由 3 条 β 片层围成一圈,螺旋直径达 0.17 ~ 0.27nm,共有 7 圈多右手螺旋。不久又在一些细菌的碱性蛋白酶及噬菌体 P₂₂尾刺蛋白(tailspike protein)中发现类似的模体结构,后者有水解沙门杆菌 O 抗原糖链的功能。

已提出的模体类型还很多,如形似丹麦环饼的三环域(Kringle domain)模体、反平行四 β 链的上皮生长因子模体、各种形状的 Ω 模体等。随着蛋白质结构的进一步阐明,似乎还有愈加增多的趋势。由于模体大小差别不一,故可属于不同的蛋白质结构水平,不能单纯以模体作为蛋白质分类的标准。然而,蛋白质模体概念的提出为描述蛋白质结构形状提出了简便而直观的思考模式,虽然至今对模体的认识尚有限,有的功能仍不清楚,有待于进一步深入研究以充实蛋白质空间结构与其功能关系的内容。孙之荣从 240 个高精度结构的蛋白质库中总结出四大类型超二级结构模体,即 α -拐角、 α -发夹、 β -发夹及拱形结构,包含着 11 种构象和氢键模式,这些简单模体还仅是复杂模体的基础。充分认识了模体的结构与功能或许有利于设计由多个不同模体连接在一条肽链上以制造具有新功能的杂交蛋白质(hybritein),以适应人们更高的要求。

4. 蛋白质的四级结构

由两条以上具有独立结构的多肽链通过非共价键相互结合而成的结构是蛋白质的四级结构(quadernary structure)。构成四级结构中的每条具有独立三级结构的多肽链称为亚基或亚单位(subunit)。蛋白质分子作为一个结构性或功能性物质的整体,有许多是由两条或更多多肽链(相同或不相同亚基)形成稳定的空间结构,它们之间借非共价键结合,这种结合实际上是亚基间的缔合,而四级结构的严格定义应是涉及这类寡聚体蛋白质中各亚基的空间排布及各亚基间的相互作用,不包