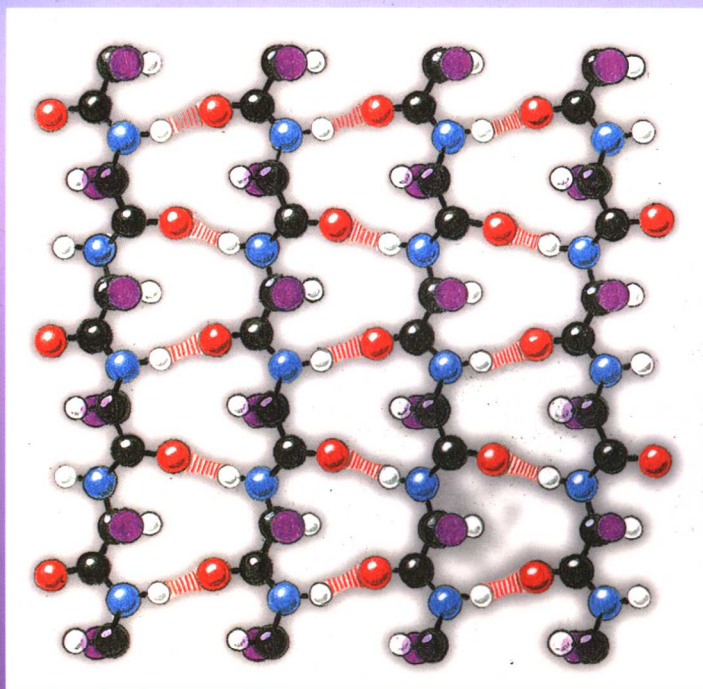


Taylor & Francis  
Taylor & Francis Group

# 蛋白质结构 导论

卡尔·布兰登 约翰·图兹 著  
王克夷 龚祖坝 译



上海科学技术出版社

# 蛋白质结构导论

[瑞典]卡尔·布兰登 [英]约翰·图兹 著  
王克夷 龚祖坝 译

上海科学技术出版社

## 图书在版编目 (C I P) 数据

蛋白质结构导论 / (瑞典) 布兰登, (英) 图兹著;  
王克夷, 龚祖坝译. — 上海: 上海科学技术出版社,  
2007.1

ISBN 978-7-5323-8377-1

I. 蛋... II. ①布... ②图... ③王... ④龚...  
III. 蛋白质-生物结构-概论 IV. Q510.1

中国版本图书馆CIP数据核字 (2006) 第018212号

## **Introduction to Protein Structure (Second Edition)** **by Carl Branden and John Tooze**

All Rights Reserved

Authorized translation from English language edition  
published by Garland, a member of the Taylor &  
Francis Group.

上海世纪出版股份有限公司 出版、发行  
上海科学技术出版社

(上海钦州南路71号 邮政编码 200235)

新华书店上海发行所经销

上海精英彩色印务有限公司印刷

开本 889 × 1194 1/16 印张 22.75

字数 720 000

2007年1月第1版 2007年1月第1次印刷

印数 1-2 000

定价: 160.00元

---

本书如有缺页、错装或坏损等严重质量问题,  
请向承印厂联系调换

# 译者序

20世纪下半叶，生物学研究，尤其是分子水平上的生物学研究，取得了很大进展。人们虽然早已了解，核酸和蛋白质是组成生命的两类最重要的分子，但是显然在前者的研究中投入了更多关注。遗传信息的阐明、个别基因序列乃至人类全基因组序列的测定，这些光彩夺目的成就不能不令人为之折服。然而，还有一个蛋白质的世界，它更加多样化，对维持生命活动起着同样重要的作用；研究者们在其中默默耕耘，投入了无数的精力。

从上世纪50年代起，人们就已认识到蛋白质结构与功能之间的关系。结构是蛋白质行使功能的基础，而功能是蛋白质结构的生化表现特征。遗传工程出现后不久，立刻就出现了蛋白质工程，其发展必然依赖于对蛋白质结构的了解。工业上的许多要求，如化学工业中对酶的应用和改造、医药工业中的药物设计，又推动着蛋白质结构研究的快速发展。

本书是卡尔·布兰登和约翰·图兹编写的《蛋白质结构导论》第二版。它不仅以深浅有序、内容翔实、涉及面广（包括介绍DNA结构和病毒结构）的特色吸引了我们的注意，而且作者在注重结构-功能关系的同时，把蛋白质分子与蛋白质分子之间、核酸分子与蛋白质分子之间的相互作用放在突出地位。本书的另一吸引人之处是有大量彩色插图，显示多种多样的分子结构，为读者理解所阐述的内容提供了莫大方便。

目前人们正期待着结构生物学研究在“后基因组时代”有更为长足的发展。基因组序列的快速测定为阐明蛋白质的一级结构打下了坚实基础，而且随着结构测定技术的日臻完善，被阐述的生物大分子的高级结构，特别是蛋白质的高级结构，可谓与日俱增。正如作者所言，1988年开始撰写本书第一版时，在中等和高分辨率上得到阐明的蛋白质高级结构约有250种左右；1999年本书第二版问世时，被阐明的蛋白质高级结构已超过6500种。从我们着手翻译到目前出版，又过了五年多，相信又有数千种已阐明的蛋白质高级结构未能纳入书中。不过，本书的整个内容还是着重于领域内的基础知识，所阐述的原则对各种类型的蛋白质分子结构仍然适用，因此可供生物学、分子生物学、蛋白质化学、结构生物学、蛋白质工程、药物和农学设计、酶学和抑制剂等领域的科研工作者、教师和学生阅读及参考。

浙江大学医学院生化教研室詹金彪教授和中国科学院上海有机化学研究所夏宗芴研究员对本书部分章节的译稿提出了宝贵意见，特此感谢。

王克夷 龚祖坝

2005年12月20日

## 第二版序言

自从1991年本书第一版问世以来，蛋白质原子结构的测定已受到来自多方面的推动。每年报道的新结构近乎成倍地增加。技术的进步历历可见，比如同步加速器的X射线更加便于使用，冷冻晶体的方法降低了射线对晶体的损伤，多维NMR以及更强磁场的NMR仪器得到发展，还有基因克隆、测序和表达系统的开发。这些都对蛋白质结构测定发展作出了贡献。一方面，较多量地获得自然界罕见的蛋白质越来越容易；另一方面，晶体学家只需较以前更小的晶体就能开展工作。

分子生物学家有一个基本信条：不了解相关分子的结构，便无法阐明生物反应。这一信条终究是站得住脚的。蛋白质结构数据库正在迅速扩大，生物学路径的范围也随之扩大。关于这些路径，我们可以在接近原子的分辨率上提出许多有意义的问题。阐明蛋白质的结构原则对于分子生物学正变得具有更广泛的意义。

在过去十年中，制药工业成了蛋白质数据库的主要用户，也成了测定新结构的主要贡献者。了解一个酶或受体的原子结构，对于探索专一而且强烈结合的抑制剂显得至关重要。例如，寻求HIV蛋白酶的有效抑制剂，用于艾滋病的综合治疗。这一探索引导着许多制药公司去测定蛋白酶结合抑制剂的结构。这样的结构测定迄今已完成了120多个，而且至少有两种HIV蛋白酶抑制剂目前被用于艾滋病的常规治疗。看来可以肯定，测定靶分子的原子结构将会在药物设计中起到越来越重要的作用。

对于蛋白质结构的不断阐明导致了商业上的开发，这类开发当然不限于制药工业。酶在化学工业中的应用、新的更专一的杀虫剂和除草剂的研制、为改变植物油和植物糖类而对酶进行的改造，这些都是其他商业开发途径的例子。它们一定程度上均有赖于在高分辨率上了解特定的蛋白质结构。

时值更多物种的全基因组被测序之际，测定以前未加鉴定的读码框之功能，正成为愈益具有挑战性的问题。人们在认真商讨建立一个自动化、高通量的结构测定中心之可能性。如果某种蛋白质在序列上与已知功能的蛋白质没有可识别的同源性，那么建立测定中心是通过结构同源性来鉴定其功能的有效办法，虽然这听起来或许有点惊人。

过去十年里，对高分辨率蛋白质结构的兴趣在增长，同时本书的第一版又颇受欢迎。这两件事鼓励我们准备出版本书的新版本。目前一些大学给有关蛋白质结构的课程投入了更多的课时，或者在更加基础的生物学和生物化学课程中增加了有关蛋白质结构的授课时数。我们希望新版的《蛋白质结构导论》将对教师和学生皆有助益。

1988年开始撰写第一版时，仅有约250个蛋白质结构的测定达到中、高分辨率，而且在那个时候，专业的蛋白质晶体学家大多彼此熟识，因此编写时未曾遇到材料严重遗漏的问题。如今，在纽约布鲁克黑文蛋白质数据库中收集的蛋白质坐标已经超过6 500个。结构的数目和相关生物体系的种类都是非常之多，因而蛋白质结构这一领域变得更加零碎、更加专门化，犹如置身于众多树木中要看到一片树林那么困难。所以对当今的作者来说，如何取舍变得极为重要。鉴于此，我们力避喋喋于多少多少的蛋白质，徒然添加琐细而不促进对基本概念的理解。新版中删去两章，同时增加四章，故而无可避免地增加了一些篇幅。我们斗胆未提及一些结构生物学家有心得的蛋白质，或许令他们失望。在此我们谨表歉意，并祈海涵。

## 第二版致谢

在编写本书第二版时，我们再次大力仰仗并得益于诸多同行的忠告和建设性评议。对Ken Holmes(马普学会海德堡研究所)、Lawrence Stern(麻省理工学院)、Michelle Arkin(Sunesis制药厂)和Watson Fuller(英国基尔大学)特表谢忱，他们分别为第14、18章，第15章，以及第17、18章有所贡献。Stephen Harrison(哈佛大学)和Paul Sigler(耶鲁大学)分别就第8~10、13和16章，以及第8~10、13章给予广泛的帮助和忠告，我们铭感尤深。

以下各位(按字母顺序)曾审阅了一章或多章，并改正书中事实或解释的讹误，或帮助确保事实与解释之间适当的平衡与重点：Tom Alber(加州大学伯克利分校)，Tom Blundell(英国剑桥大学)，Stephen Burley(洛克菲勒大学)，Charles Craik(加州大学旧金山分校)，Ken Dill(加州大学旧金山分校)，Chris Dobson(英国牛津大学)，Robert Fletterick(加州大学旧金山分校)，Richard Henderson(英国分子生物学实验室)，Werner Kuehlbrandt(马普学会法兰克福研究所)，David Parry(新西兰梅西大学)，Greg Petsko(布兰德斯大学)，David Trentham(英国伦敦国立医学研究院)。

本书不至过分艰涩，有赖于它的插图。在这件事上，Nigel Orme功不可没。他像对待第一版那样，将草图变成明了易懂的示图。Keith Roberts再次指教我们，如何最生动地用图来表示化学的和结构的现象。Jane Richardson(杜克大学)慷慨惠予本书以Kinemag补充版，而全书都要靠Richardson方式的图像，才可能让讨论到的结构成为读者可理解的。还须感谢我们的出版商Garland出版社(现已成为Taylor & Francis出版集团之一部分)所给予的支持，特别是Matthew Day，他热情地编辑了全部手稿。Miranda Robertson则以她无与伦比的干练，再次安排了整个出版计划。

# 目 录

## 第一部分 基本的结构原则

第1章 建筑构件	3	四螺旋捆是 $\alpha$ 类型蛋白质中常见的域结构	30
蛋白质是多肽链	4	$\alpha$ 螺旋结构域有时大而复杂	32
遗传密码规定了20种氨基酸的侧链	4	珠蛋白折叠模式出现在肌红蛋白和血	
半胱氨酸可形成二硫键	5	红蛋白中	33
肽单元是蛋白质结构的建筑构件	5	几何因素决定 $\alpha$ 螺旋的排列	33
甘氨酸可以适应许多不同的构象	5	一段 $\alpha$ 螺旋的脊与邻近螺旋的沟相配	33
某些侧链构象是能量有利的	9	珠蛋白折叠模式在进化过程中被保留下来	34
许多蛋白质含固有的金属原子	9	疏水的内部被保留	34
结论	11	螺旋的运动适应内部侧链的突变	35
第2章 蛋白质结构的模体	12	镰状细胞血红蛋白产生对疟疾的抗性	35
蛋白质的内部是疏水的	12	结论	37
$\alpha$ 螺旋是一种重要的二级结构元件	13	第4章 $\alpha/\beta$ 结构域的结构	38
$\alpha$ 螺旋有一个偶极矩	14	平行的 $\beta$ 折叠链被安排在桶或片层中	38
一些氨基酸倾向于 $\alpha$ 螺旋	14	$\alpha/\beta$ 桶存在于许多不同类型的酶中	39
$\beta$ 片层通常含有平行或反平行的 $\beta$ 折叠		分支的疏水侧链在 $\alpha/\beta$ 桶的核心中占	
链	15	主导地位	40
环区位于蛋白质分子的表面	17	丙酮酸激酶含有几个结构域,其中之一	
蛋白质的立体结构图突出二级结构	19	是 $\alpha/\beta$ 桶	41
拓扑图对蛋白质结构的分类有用	19	通过基因融合形成双桶结构	41
二级结构元件被连接,形成简单的模体	20	位于 $\alpha/\beta$ 桶一端的环区构成活性部位	42
发夹 $\beta$ 模体经常出现在蛋白质结构中	23	$\alpha/\beta$ 桶为新酶活性的进化提供了例证	43
希腊钥匙模体被发现于反平行 $\beta$ 片层中	23	富亮氨酸模体形成 $\alpha/\beta$ 马蹄形折叠模式	44
$\beta$ - $\alpha$ - $\beta$ 模体含有两股平行 $\beta$ 折叠链	24	$\alpha/\beta$ 扭曲的敞开片层结构在 $\beta$ 片层的两	
蛋白质分子被组织在一定的层次结构中	25	侧含有 $\alpha$ 螺旋	45
大的多肽链被折叠到几个结构域中	25	敞开的 $\beta$ 片层结构具有多种拓扑学	45
结构域由结构模体构成	25	在 $\alpha/\beta$ 结构中能预测活性部位的位置	45
一些简单的模体组合形成复杂的模体	26	酪氨酰-tRNA合成酶有两个不同的结构	
蛋白质结构可以分为三种主要的类型	26	域( $\alpha/\beta+\alpha$ )	46
结论	27	羧肽酶是一种带混合 $\beta$ 片层的 $\alpha/\beta$	
第3章 $\alpha$ 结构域的结构	29	蛋白质	49
卷曲之卷曲的 $\alpha$ 螺旋含有重复的七氨基		阿拉伯糖结合蛋白有两个相似的 $\alpha/\beta$	
酸序列模式	29	结构域	50
		结论	51

**第5章  $\beta$ 结构** .....52

  上下走向的桶具有已知的简单拓扑学 .....52

  视黄醇结合蛋白在一个上下走向的桶内  
    结合视黄醇 .....52

  氨基酸的序列反映了 $\beta$ 结构 .....53

  视黄醇结合蛋白属于一个蛋白质结构的  
    超家族 .....54

  神经氨酸酶折叠为上下走向的 $\beta$ 片层 .....54

  折叠模体在神经氨酸酶中形成一个螺旋  
    桨样的结构 .....55

  活性部位在螺旋桨一侧的中间 .....56

  希腊钥匙模体往往存在于反平行 $\beta$ 结构中 .....57

$\gamma$ 晶体蛋白分子有两个结构域 .....57

  这个域结构具有简单的拓扑学 .....58

  两个希腊钥匙模体构成这个结构域 .....58

  两个结构域具有相同的拓扑学 .....58

  两个结构域具有相似的结构 .....59

$\gamma$ 晶体蛋白中的希腊钥匙模体在进化上  
    有关联 .....59

  希腊钥匙模体可形成胶冻卷样的桶 .....60

  胶冻卷模体围裹在一个桶的周围 .....60

  胶冻卷样桶往往可分为两个片层 .....61

  有功能的血凝素亚基含有两条多肽链 .....61

  亚基的结构分为茎部和顶部 .....61

  受体的结合部位由胶冻卷结构域构成 .....62

  血凝素起膜融合原的作用 .....63

  血凝素的结构因pH的改变而受到影响 .....64

  平行 $\beta$ 螺旋结构域有一种新的折叠模式 .....65

  结论 .....68

**第6章 折叠和柔性** .....69

  球状蛋白质仅处于稳定状态的边缘 .....69

  动力学因素对折叠过程是重要的 .....70

  熔球态是折叠过程的中间态 .....71

  疏水侧链的包埋是一个关键的事件 .....71

  单一的和多重的折叠途径都已被观察到 .....72

  在折叠过程中酶帮助形成合适的二硫键 .....73

  脯氨酸残基的异构化可能是蛋白质折叠  
    时限制速率的步骤 .....75

  蛋白质可在伴侣蛋白内进行折叠和去  
    折叠 .....77

  GroEL是一个带有中心通道的圆柱结构,  
    新合成的多肽链结合于通道中 .....77

  GroES关闭GroEL圆柱体的一端 .....79

  GroEL-GroES复合物在一个依赖ATP  
    的循环中结合和释放新合成的肽链 .....80

  折叠的状态仍具有可变的结构 .....81

  在一种蛋白激酶中, 构象的改变对调节  
    细胞周期是重要的 .....81

  钙调蛋白与肽的结合可诱导内部结构域  
    的一个大的运动 .....85

  抑丝酶通过一个装载弹簧的安全捕获  
    机制抑制丝氨酸蛋白酶 .....86

  效应分子使别构蛋白质在R状态和T状  
    态之间转换 .....88

  由X射线衍射研究所得的结构, 可解释  
    磷酸果糖激酶的别构特性 .....88

  结论 .....92

**第7章 DNA的结构** .....93

  DNA双螺旋在A-DNA和B-DNA中是  
    不同的 .....93

  DNA螺旋有大沟和小沟 .....93

  Z-DNA形成锯齿状图形 .....94

  B-DNA在体内是优先构象 .....94

  B-DNA中专一的碱基序列可被识别 .....95

  结论 .....97

**第二部分 结构、功能和工程**

**第8章 原核生物通过螺旋-转角-螺旋  
模体的DNA识别** .....101

  基因调控的分子机制 .....101

  阻遏蛋白和Cro蛋白操纵原核生物的基  
    因开关区 .....101

  完整 $\lambda$ 噬菌体Cro蛋白的X射线衍射结构  
    已被阐明 .....102

$\lambda$ 阻遏蛋白的DNA结合结构域的X射线  
    衍射结构已被阐明 .....103

$\lambda$ 噬菌体Cro蛋白和阻遏蛋白均有一个  
    专一的DNA结合模体 .....105

  由模型构建预测Cro-DNA相互作用 .....105

  基因研究与这一结构模型相合 .....105

  DNA与噬菌体434Cro及阻遏蛋白之



复合物的X射线衍射结构揭示出蛋白质-DNA相互作用的一些新特征 .....	106	了解致肿瘤突变 .....	128
噬菌体434Cro的结构颇类似于434阻遏蛋白DNA结合结构域的结构 .....	108	单体p53多肽链可分为三个结构域 .....	128
蛋白质使复合物中的B-DNA发生细微畸变 .....	108	寡聚结构域形成四聚体 .....	129
通过序列专一的蛋白质-DNA相互作用识别操纵基因区 .....	109	p53的DNA结合结构域是一个反平行 $\beta$ 桶 .....	129
蛋白质-DNA主链骨架的相互作用决定DNA构象 .....	110	p53 的两个环区和一个 $\alpha$ 螺旋结合DNA .....	131
DNA构象变化对阻遏蛋白和Cro有差别地结合于不同的操纵基因位点是重要的 .....	111	致肿瘤突变主要发生在参与DNA结合的三个区域 .....	131
噬菌体阻遏蛋白和Cro的实质 .....	111	结论 .....	133
DNA结合是由别构控制所调节 .....	111	<b>第10章 专一转录因子属于几个家族 .....</b>	<b>134</b>
trp 阻遏蛋白形成一个螺旋-转角-螺旋模体 .....	112	几类不同的含锌模体已被观察 .....	134
构象变化开启功能开关 .....	112	经典锌指沿DNA大沟串联地结合于DNA .....	135
Lac阻遏蛋白能结合大沟和小沟,引起DNA的急剧弯曲 .....	113	经典锌指模体的指区与DNA相互作用 .....	136
CAP介导的DNA弯曲能够激活转录 .....	114	糖皮质激素受体中的两个含锌模体构成一个DNA结合结构域 .....	137
结论 .....	116	糖皮质激素受体的二聚体结合于DNA .....	139
<b>第9章 真核生物转录因子对DNA的识别 .....</b>	<b>118</b>	第一个锌模体中的 $\alpha$ 螺旋提供了专一的蛋白质-DNA相互作用 .....	140
转录由蛋白质-蛋白质相互作用来激活 .....	119	识别螺旋中的三个残基提供了与DNA序列的专一相互作用 .....	141
TATA盒结合蛋白是普遍存在的 .....	119	顺式视黄酸受体形成异二聚体,能识别带有可变间隔区的串联重复区 .....	141
TBP-TATA盒复合物的三维空间结构已知晓 .....	120	酵母转录因子GAL4在其DNA结合结构域中含有双核锌簇 .....	142
TBP中的 $\beta$ 片层构成DNA结合位点 .....	121	GAL4的锌簇区结合于增强子元件的两端 .....	142
TBP结合于DNA螺旋的小沟中,并引起DNA很大的结构变化 .....	121	接头区也对DNA结合起作用 .....	143
TBP与TATA盒之间的作用区主要是疏水性的 .....	121	转录因子C <sub>2</sub> 锌簇家族内的DNA结合专一性是由接头区实现的 .....	144
由TBP引起DNA畸变的功能意义 .....	123	含锌转录因子家族以几种不同的方式结合于DNA .....	145
TF II A和TF II B结合于TBP和DNA .....	123	亮氨酸拉链为一些真核转录因子提供了二聚化的相互作用 .....	146
同源域蛋白质参与了很多真核生物体的发育 .....	123	GCN4碱性区亮氨酸拉链作为两个连续 $\alpha$ 螺旋的二聚体与DNA结合 .....	147
同源域蛋白质的单体通过一个螺旋-转角-螺旋模体与DNA结合 .....	124	GCN4通过专一的和非专一的接触与DNA结合 .....	148
同源域转录因子在体内的专一性取决于跟其他蛋白质的相互作用 .....	125	HLH模体参与同二聚体和异二聚体的缔合 .....	149
两个串联定向的螺旋-转角-螺旋模体使POU区结合DNA .....	126	b/HLH模体的 $\alpha$ 螺旋碱性区结合于DNA的大沟之中 .....	150
同源域在体内的功能尚有很多值得研究之处 .....	127	转录因子b/HLH/zip家族具有HLH和亮氨酸拉链这两种二聚化模体 .....	152
		Max和MyoD通过不同专一性的蛋白质-DNA .....	

相互作用来识别DNA HLH的共有序列 . . . . .	153
结论 . . . . .	153
<b>第11章 酶催化的一个例子：丝氨酸蛋白酶</b> . . . . .	155
蛋白酶有四个功能类型的家族 . . . . .	155
酶的催化特性反映在 $K_m$ 和 $k_{cat}$ 这两个数值中 . . . . .	155
酶降低化学反应的活化能 . . . . .	156
丝氨酸蛋白酶通过形成四面体过渡态来断裂肽键 . . . . .	156
丝氨酸蛋白酶的催化作用需要四个重要的结构特征 . . . . .	157
趋同进化产生了两种具有相似催化机制的不同丝氨酸蛋白酶 . . . . .	158
胰凝乳蛋白酶具有两个反平行 $\beta$ 桶结构域 . . . . .	159
每个结构域的两个环区构成活性部位 . . . . .	159
胰凝乳蛋白酶分子是通过基因复制进化而来的吗? . . . . .	161
优先断裂的特性取决于底物专一性口袋中的不同侧链 . . . . .	161
底物专一性口袋的工程化突变可改变催化速率 . . . . .	161
胰蛋白酶的Asp189-Lys突变体给底物专一性造成想象不到的改变 . . . . .	163
丝氨酸蛋白酶——枯草杆菌蛋白酶的构属于 $\alpha/\beta$ 类型 . . . . .	163
枯草杆菌蛋白酶和胰凝乳蛋白酶的活性部位是相似的 . . . . .	163
枯草杆菌蛋白酶的结构异常有其功能意义 . . . . .	164
利用蛋白质工程方法剖析枯草杆菌蛋白酶过渡态的稳定性 . . . . .	165
催化可发生在无催化三元组合的情况下 . . . . .	165
在底物协助的催化中，底物分子提供催化基因 . . . . .	165
结论 . . . . .	166
<b>第12章 膜蛋白</b> . . . . .	167
膜蛋白很难结晶 . . . . .	167
新的结晶方法正在被开发 . . . . .	168
膜蛋白的二维晶体可用电子显微镜进行研究 . . . . .	169
细菌视紫红质含有七段穿膜 $\alpha$ 螺旋 . . . . .	169
细菌视紫红质是一种光驱动的质子泵 . . . . .	170
孔蛋白利用 $\beta$ 折叠链形成穿膜通道 . . . . .	172
孔蛋白通道由上下走向的 $\beta$ 桶构成 . . . . .	172
每个孔蛋白分子有三个通道 . . . . .	173
离子通道将离子选择性与高水平的离子传导相结合 . . . . .	173
$K^+$ 通道是一个四聚体分子，它在四个亚基之间的界面中有一个离子孔 . . . . .	174
离子孔具有一个窄的离子选择性滤器 . . . . .	174
细菌光合反应中心由四条不同的多肽链和许多色素构成 . . . . .	175
L亚基、M亚基和H亚基具有穿膜 $\alpha$ 螺旋 . . . . .	176
光合色素结合到L亚基和M亚基上 . . . . .	177
反应中心因电子穿越膜而将光能转化为电能 . . . . .	178
天线色素蛋白装配成多聚的集光颗粒 . . . . .	179
叶绿素分子在集光复合物LH2中形成圆环 . . . . .	179
反应中心被集光复合物LH1的16个天线蛋白构成的环所包围 . . . . .	181
由氨基酸的序列能预测穿膜 $\alpha$ 螺旋 . . . . .	182
疏水性标尺可衡量不同氨基酸侧链的疏水性程度 . . . . .	182
水性作图可鉴定穿膜螺旋 . . . . .	183
反应中心的水性作图与晶体结构数据相一致 . . . . .	183
膜脂质和蛋白质的穿膜 $\alpha$ 螺旋没有专一相互作用 . . . . .	184
结论 . . . . .	184
<b>第13章 信号转导</b> . . . . .	186
G蛋白是分子放大器 . . . . .	186
Ras蛋白和 $G_q$ 催化结构域具有类似的三维结构 . . . . .	188
$G_q$ 通过三个开关区的构象改变而被激活 . . . . .	190
GTP酶通过一个水分子的亲核攻击来水解GTP . . . . .	192
$G_{\beta}$ 亚基有一个七叶片螺旋桨样的折叠模式，由七个WD重复片段构成 . . . . .	193
$G_{\alpha}$ 的GTP酶结构域在异三聚体的 $G_{\alpha\beta\gamma}$ 复合物中与 $G_{\beta}$ 结合 . . . . .	195
光导蛋白调节视杆细胞的光适应性 . . . . .	196
光导蛋白结合到 $G_{\beta\gamma}$ 上，可阻断与 $G_{\alpha}$ 的结合 . . . . .	196
人生长激素诱导共同源受体的二聚化 . . . . .	197

生长激素受体的二聚化是一个序列性的过程 .....	199	超变区紧密缔合而形成的 .....	226
生长激素也结合催乳激素受体 .....	199	抗原结合部位在沟槽中结合半抗原, 在大而平坦的表面结合蛋白质抗原 .....	227
酪氨酸激酶受体是酶连接的重要受体 .....	201	除重链CDR3以外, CDR环仅承担有限种类的构象 .....	230
小的蛋白质模件构成信号转导网络的衔接者 .....	202	一种IgG分子有多种程度的构象灵活性 .....	231
SH2结构域结合到靶分子中含磷酸化酪氨酸的区域 .....	202	MHC分子的结构已经为了解T细胞激活的分子机制提供了思路 .....	231
SH3结构域与靶分子中富含脯氨酸的区域相结合 .....	203	MHC分子由抗原结合结构域和免疫球蛋白样结构域组成 .....	232
Src酪氨酸激酶除含有酪氨酸激酶区外, 还含有SH2结构域和SH3结构域 .....	205	抗体识别在MHC分子中不同于在免疫球蛋白中 .....	233
失活状态激酶的两个结构域因调节结构域的装配而处于一个紧密的构象中 .....	206	肽链以不同的方式被 I 型和 II 型MHC分子所结合 .....	234
结论 .....	207	T细胞受体具有可变和不变的免疫球蛋白结构域, 还具有超变区 .....	234
<b>第 14 章 纤维状蛋白质</b> .....	209	MHC-肽复合物是T细胞受体的配体 .....	235
胶原是一种由三股平行、十分伸展的左手螺旋所构成的超螺旋 .....	209	许多细胞表面受体含有免疫球蛋白样结构域 .....	236
卷曲之卷曲常用于形成纤维状蛋白质和球状蛋白质的多聚体 .....	211	结论 .....	237
淀粉样纤维被认为是由连续的 $\beta$ 片层螺旋所构成 .....	212	<b>第 16 章 球状病毒的结构</b> .....	239
蜘蛛丝是自然界中最高性能的纤维 .....	214	球状病毒的蛋白质外壳具有二十面体对称 .....	240
肌纤维含有肌球蛋白和肌动蛋白, 肌肉收缩时两者彼此相对滑动 .....	215	二十面体具有高度的对称性 .....	240
肌球蛋白头部构成肌动蛋白与肌球蛋白丝之间的联桥 .....	215	最简单的病毒具有60个蛋白质亚基组成的外壳 .....	241
对蛙肌的时间分辨X射线衍射研究证明了联桥的移动 .....	216	复合球状病毒在非对称单元内有不止一条多肽链 .....	242
肌动蛋白和肌球蛋白的结构已被测定 .....	216	在T=3的植物病毒中, 结构的多面性提供了准等同排列 .....	243
肌球蛋白的结构支持摆动联桥假设 .....	218	蛋白质亚基能识别在外壳内部的RNA专一部位 .....	244
肌肉收缩中ATP的功能类似于G蛋白活化中GTP的功能 .....	219	小RNA病毒的蛋白质外壳含有四条多肽链 .....	244
结论 .....	220	小RNA病毒有四种不同的结构蛋白质 .....	245
<b>第 15 章 免疫系统对外来分子的识别</b> .....	221	小RNA病毒外壳内的亚基排列类似于T=3的植物病毒 .....	245
抗体多肽链分为若干结构域 .....	222	许多不同的球状植物病毒和动物病毒的外壳蛋白质具有相似的胶冻卷桶形结构, 表明它们的进化关系 .....	246
抗体的多样性产生于几种不同的机制 .....	223	由鼻病毒的结构可以设计针对感冒病毒的药物 .....	248
所有的免疫球蛋白具有类似的三维结构 .....	224	噬菌体MS2具有不同的亚基结构 .....	248
免疫球蛋白折叠模式被最清楚地描述为两个彼此紧密排列的反平行 $\beta$ 片层 .....	225		
超变区簇聚在可变结构域一端的环区中 .....	226		
抗原结合部位是通过重链和轻链两者的			

MS2亚基的二聚体能识别RNA包装信号 . . . . .	250	仅改变序列中的半数残基, 一个 $\beta$ 结构 即可转换为 $\alpha$ 结构 . . . . .	270
$\alpha$ 病毒的核心蛋白质具有类似于胰凝乳蛋 白酶的折叠模式 . . . . .	250	结论 . . . . .	271
SV40和多瘤病毒的外壳由主要外壳蛋白质 的五聚体构成, 它们是不等构装配的, 但大体上又是等构相互作用的 . . . . .	251	<b>第 18 章 蛋白质结构的测定</b> . . . . .	272
结论 . . . . .	253	几种不同的技术用于研究蛋白质分子 的结构 . . . . .	272
<b>第 17 章 蛋白质结构的预测、工程化和设计</b> . . . . .	255	蛋白质晶体很难生长 . . . . .	272
同源蛋白质具有类似的结构和功能 . . . . .	255	X射线源可以是单色的, 也可以是多 色的 . . . . .	274
同源蛋白质具有保守的结构核心和可变 的环区 . . . . .	256	X射线衍射数据或是记录在图像板上, 或是用电子检测仪记录 . . . . .	275
二级结构的知识对预测三级结构是必 需的 . . . . .	256	Bragg 定律给出了衍射的规则 . . . . .	275
预测二级结构的方法可得益于同源蛋白 质的多重排比 . . . . .	257	相位测定是晶体学中的主要问题 . . . . .	276
许多不同的氨基酸序列给出类似的三维 结构 . . . . .	258	通过多波长的异常衍射实验, 也可得 到相位信息 . . . . .	278
从序列预测蛋白质结构是一个尚未解 决的问题 . . . . .	258	模型之构建包含了对数据的主观解释 . . . . .	278
穿引法可将一些氨基酸序列归属于一些 已知的三维折叠模式 . . . . .	259	通过精修可以去除初始模型中的错误 . . . . .	279
通过工程化可使蛋白质更为稳定 . . . . .	259	近年在技术上的进展对蛋白质晶体学 产生了重大影响 . . . . .	280
二硫桥键增加蛋白质的稳定性 . . . . .	260	X射线衍射可用于研究纤维及晶体的 结构 . . . . .	280
甘氨酸和脯氨酸对稳定性起相反的作用 . . . . .	261	生物高聚物的结构可以用纤维衍射的 方法加以研究 . . . . .	281
稳定 $\alpha$ 螺旋的偶极矩可增加稳定性. . . . .	261	核磁共振(NMR)方法利用原子核的磁学 特性 . . . . .	282
一些使疏水核心之空穴填满的突变 体, 并不能使T4溶菌酶更稳定 . . . . .	262	用序列归属方法解释蛋白质的二维NMR 谱 . . . . .	283
利用组合方法可设计蛋白质 . . . . .	262	距离约束被用于推算蛋白质分子的可能 结构 . . . . .	285
噬菌体展示法把蛋白质库与DNA相连接 . . . . .	262	生化研究和分子结构给出互补的功能 信息 . . . . .	286
经噬菌体展示, 蛋白酶抑制剂的亲和性 和专一性可以最优化 . . . . .	264	结论 . . . . .	286
结构支架可以缩小, 而功能不变 . . . . .	265	<b>万维网上的蛋白质结构信息</b> . . . . .	289
随机肽库的噬菌体展示可鉴定红细胞 生成素受体的激动剂 . . . . .	266	<b>参考文献</b> . . . . .	291
DNA混组可加速基因的进化 . . . . .	267	<b>索引</b> . . . . .	313
蛋白质结构可按照最基本的原理进行设计. . . . .	269		

# 第一部分

## 基本的结构原则



# 第1章 建筑构件

重组DNA技术已经提供了快速测定DNA序列的工具,而且可以由结构基因推断蛋白质的氨基酸序列。这些序列的数量目前几乎呈指数式增长,但是序列本身所能告诉我们的有关系统的生物学知识极少,并不比纽约市电话簿所展示的城市功能和有趣事物更多。

人们在自然界中观察到的蛋白质因选择性的压力而进化,以至于能实施诸多专一的功能。蛋白质的功能特性依赖于它们的三维结构。三维结构的产生是因为多肽链中特定的氨基酸序列经折叠而由线性链构成有特异三维结构的紧密的结构域(图1.1)。折叠的结构域可以作为模块来构建更大的组装体,例如病毒颗粒或肌纤维;或者提供诸如酶、输氧蛋白质、调节DNA功能的蛋白质等的专一性催化部位或结合部位。

为阐明蛋白质的生物功能,人们希望能从氨基酸序列推断或预测三维结构,然而目前还不能。尽管在过去25年中付出了很大的努力,折叠问题仍未解决,并且继续作为分子生物学中一个最基本的问题,向人

类的智力提出挑战。

蛋白质折叠依然是一个问题,因为20种不同的氨基酸可以组合成许多不同的蛋白质。这个数字大于已知宇宙中的原子数目。此外还有许多方法,可以通过不同的氨基酸序列使蛋白质产生类似的结构域。与此相比,仅有四种不同核苷酸作为结构单元的DNA结构则相对简单、规则并可以预测。

蛋白质的三维结构不能预测,只能用X射线晶体学、电子晶体学或核磁共振(NMR)等技术,通过实验来加以测定。在过去30年中,6000多种蛋白质的结构已经搞清,而被测序的蛋白质则超过500000个。这样就形成了一个庞大的信息库,从中衍生出一组蛋白质结构原则。这些原则使人们易于了解蛋白质结构是如何产生的,鉴定共同结构的主要特征,联系结构和功能之间的关系,并观察不同蛋白质间的基本关系。有关蛋白质结构的科学目前正处于分类学的水平,人们开始能够从已知三维结构的少量蛋白质中识别出为数不多的一些模式和模体(motif)。

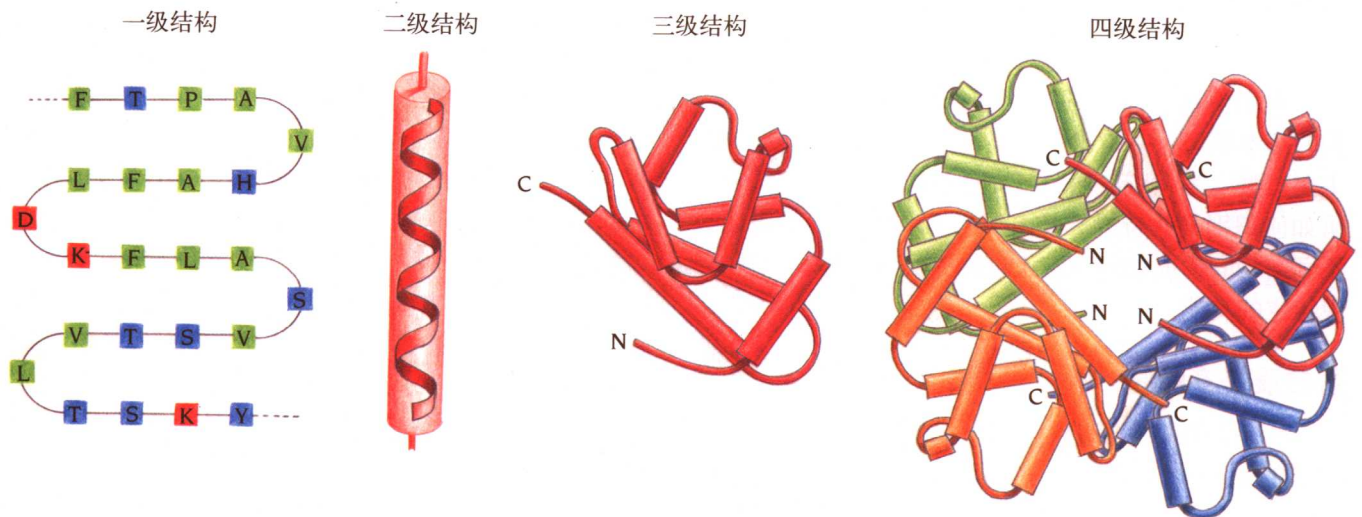


图1.1 蛋白质多肽链的氨基酸序列称为蛋白质的一级(primary)结构。序列的不同区域形成局部有规律的二级(secondary)结构,如 $\alpha$ 螺旋或 $\beta$ 折叠链。这样的结构被排列为一个或数个称为结构域的紧密球状单元,从而形成三级

(tertiary)结构。最终的蛋白质还可能以四级(quaternary)结构的方式含有数条多肽链。通过形成三级结构和四级结构,在序列上相距很远的氨基酸可以在三维结构中被紧密地安排在一起,从而形成一个功能区,即活性部位(active site)。

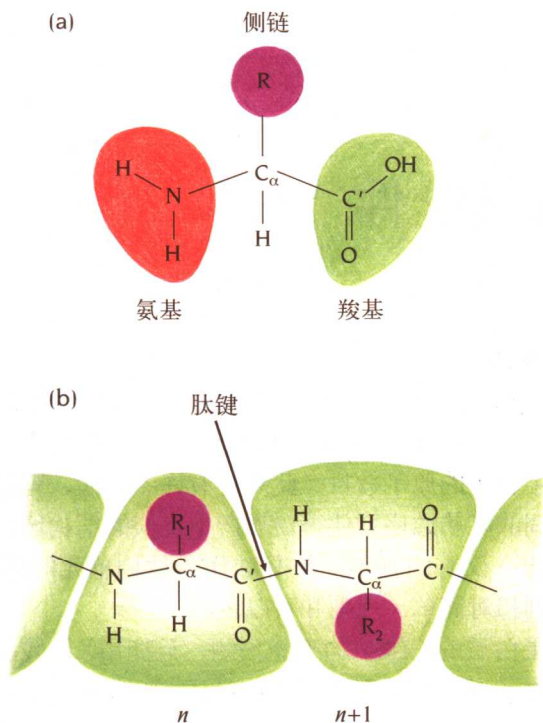


图1.2 蛋白质由氨基酸组成。氨基酸通过肽键连接为肽链。(a)一个氨基酸的图解,用以说明正文中所用的定义。一个中心碳原子( $C_{\alpha}$ ),与之相连的是一个氨基( $NH_2$ )、一个羧基( $COOH$ )、一个氢原子( $H$ )和一个侧链( $R$ )。(b)多肽链中第 $n$ 个氨基酸的羧基与第 $n+1$ 个氨基酸的氨基形成肽键,在此过程中失去一个水分子。名为残基的重复单元分为主链原子和侧链。在所有残基中,参与主链的部分是相同的,含一个中心 $C_{\alpha}$ 碳原子,以及与之相连的一个 $NH$ 基团、一个 $C'=O$ 基团和一个氢原子。侧链 $R$ 对不同的残基是不同的,它也连接在 $C_{\alpha}$ 碳原子上。

本书的前六章论述人们迄今已了解的蛋白质结构的基本原理,并给出一些实例来介绍蛋白质结构的主要类型。第7章简要讨论DNA的结构,着重于蛋白质如何识别某些特异的核苷酸序列。其余各章说明,不同结构的形成如何在进化过程中受到选择,以便实现特定的功能。

## 蛋白质是多肽链

所有20种氨基酸都有一个中心碳原子( $C_{\alpha}$ ),与之相连的是一个氢原子、一个氨基( $NH_2$ )和一个羧基( $COOH$ )(图1.2a)。用于区分不同氨基酸的是通过第四个价键与 $C_{\alpha}$ 连接的侧链。有20种侧链被基因密码所规定,其他一些偶尔存在的侧链是在翻译后经酶修饰的产物。

在蛋白质合成过程中,氨基酸的末端与末端之

间通过形成肽键(peptide bonds)而相互连接。此时,一个氨基酸的羧基与下一个氨基酸的氨基缩合,并失去一份水(图1.2b)。链伸长时,此过程被重复。其结果为,多肽链第一个氨基酸的氨基和最后一个氨基酸的羧基是完整的,而且据说链的伸展是从氨基端到羧基端。连续形成的肽键产生了“主链”,或曰“骨架”,各种侧链从主链向外伸出。

主链的原子是 $C_{\alpha}$ 碳原子,以及与之相连的 $NH$ 基团上的 $N$ 原子和羰基 $C'=O$ 上的 $C'$ 碳原子。这些单元也称为残基,通过一个残基的 $C'$ 和下一个残基的 $N$ 原子之间的肽键连接成为多肽(见图1.2b)。从生化或遗传学的观点看,沿着主链的重复结构单元就是如此( $NH-C_{\alpha}H-C'=O$ ),即这些重复结构单元是在肽键形成后氨基酸残基的共同部分(图1.2b)。

## 遗传密码规定了20种氨基酸的侧链

存在于蛋白质中的20种不同的侧链被列于图版1.1中。它们的名称以及三字母和单字母的缩写也见于该图版。单字母缩写尤其值得记住,因为它们文献中被广泛使用。一种将单字母缩写与氨基酸名称相联系的记忆法也在图版1.1中给出。

根据侧链的化学性质,氨基酸通常分为三个不同的类型。第一个类型包括那些带有严格疏水侧链的氨基酸:丙氨酸(Ala, A)、缬氨酸(Val, V)、亮氨酸(Leu, L)、异亮氨酸(Ile, I)、苯丙氨酸(Phe, F)、脯氨酸(Pro, P)和甲硫氨酸(Met, M)。四个带电的氨基酸为天冬氨酸(Asp, D)、谷氨酸(Glu, E)、赖氨酸(Lys, K)和精氨酸(Arg, R),它们构成第二类。第三类包括其他带极性侧链的氨基酸:丝氨酸(Ser, S)、苏氨酸(Thr, T)、半胱氨酸(Cys, C)、天冬酰胺(Asn, N)、谷氨酰胺(Gln, Q)、组氨酸(His, H)、酪氨酸(Tyr, Y)和色氨酸(Try, W)。甘氨酸(Gly, G)只有一个氢原子的侧链,因此是20种氨基酸中最简单的,通常或分属于第四类,或分属于第一类。

除了甘氨酸有两个氢原子结合 $C_{\alpha}$ 碳原子以外,所有氨基酸中连接 $C_{\alpha}$ 碳原子的四个基团均具有不同的化学性质。因此,甘氨酸以外的所有氨基酸都是手性分子,能够以两种不同的“手”性形式存在,即L型和D型(图1.3)。

生物系统依赖于对手性有差别的分子专一和精细的识别。蛋白质合成的翻译机器已进化到只能利用氨基酸手性中的一种。所有存在于蛋白质中的氨基酸都是L型。然而还没有明显的理由可解释,为何在进化中选用的是L型而非D型。



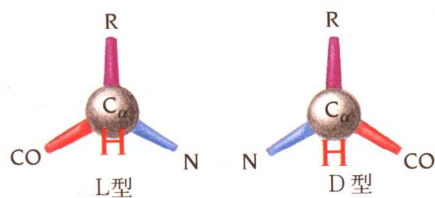
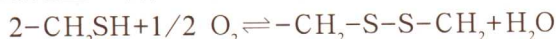


图1.3 氨基酸的手性。从氢原子俯视H-C<sub>α</sub>键，L型有来自C<sub>α</sub>的取代基CO、R和N，是顺时针的。有一个记忆的方法，即对于L型按顺时针读出CORN。

### 半胱氨酸可形成二硫键

位于多肽链不同部分、但在蛋白质三维结构上邻近的两个半胱氨酸，可被氧化而形成二硫桥键(disulfide bridge)(图1.4)。根据下列反应式，二硫桥键通常是空气氧化的最终产物：



此反应需要一个氧化环境，而且这样的二硫桥键一般未见于细胞内的蛋白质中，细胞内的蛋白质基本上是在还原的环境中“度过一生”的。然而，二硫桥键确实存在于细胞所分泌的细胞外蛋白质中，并且在真核细胞中，这些键形成于分泌途径的第一个区室

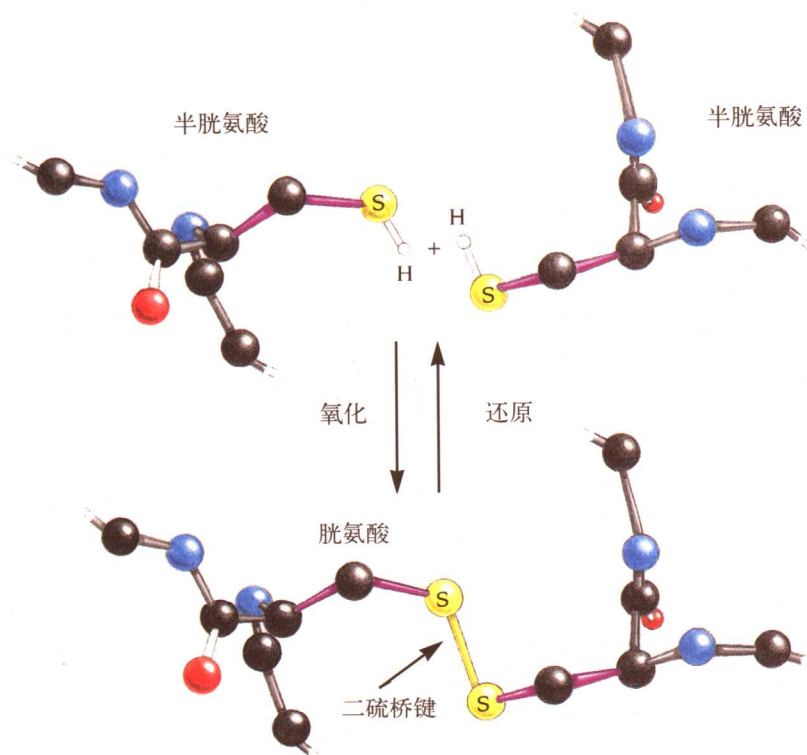
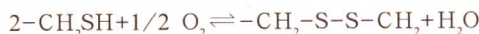


图1.4 二硫桥键通常是空气氧化的最终产物，按照下列反应式：



二硫桥键形成于两个半胱氨酸的侧链之间。来自半胱氨酸残基的两个SH基团可以在氨基酸序列的不同部分，但在三维结构中却是相邻的。它们可被氧化，形成S-S(二硫)基团。

——内质网的腔内。

二硫桥键稳定了三维结构。在有些蛋白质中这些桥键将不同肽链维系在一起，例如胰岛素的A链和B链被链间的两对二硫桥键连接。更常见的是，分子内的二硫桥键稳定了单一多肽链的折叠，致使蛋白质不易被降解。有许多这样的例子，诸如蛇毒素蛋白和蛋白酶抑制剂等由短的肽链所构成的蛋白质。当然，以上蛋白质还需要其他的稳定因素才能产生稳定的折叠模式。当前，人们致力于通过定点突变的方法，在一些酶分子中引入额外的分子内二硫桥键，目的是使酶更为热稳定，从而在工业应用中更有价值，如第17章所述。

### 肽单元是蛋白质结构的建筑构件

图1.2显示了分割多肽链的一种方法，这是生物化学家的方法。然而还有一种把主链分割为重复单元的方法，更适于描述蛋白质的结构特性。为上述目的，更有用的是将多肽链分割为从一个C<sub>α</sub>碳原子到下一个C<sub>α</sub>碳原子的肽单元(见图1.5)。这样，除了第一个和最后一个以外，其他每一个C<sub>α</sub>碳原子均属于两个这样的单元。以这种方法分割肽链的理由是，在此单元中，所有原子都被固定在一个平面内，所有蛋白质中所有这种平面上的键长和键角几乎是相同的。请注意，主链的肽单元并不包含多种不同的侧链(图1.5)。本书将同时使用两种描述多肽链的方法——生化的和结构的，而且根据不同氨基酸的序列和平面肽单元的序列，来讨论蛋白质。

通过C<sub>α</sub>碳原子上的共价键连接进入主链的肽单元是有效的刚性基团，它们被许可的自由度只是绕着两个键旋转：C<sub>α</sub>-C'和N-C<sub>α</sub>(图1.6)。习惯上，绕N-C<sub>α</sub>键旋转的角被称为phi(φ)，绕着同一C<sub>α</sub>原子的C<sub>α</sub>-C'键旋转的角被称为psi(ψ)。

在此方式中，每一氨基酸残基与两个构象角φ和ψ有密切关系。因为它们是仅有的两个自由度，所以当每个氨基酸的φ和ψ被精确地确定后，多肽的整个主链的构象就被完全决定了。

### 甘氨酸可以适应许多不同的构象 9

对一个氨基酸而言，大多数的φ和