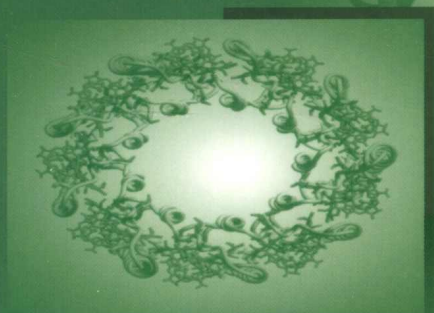
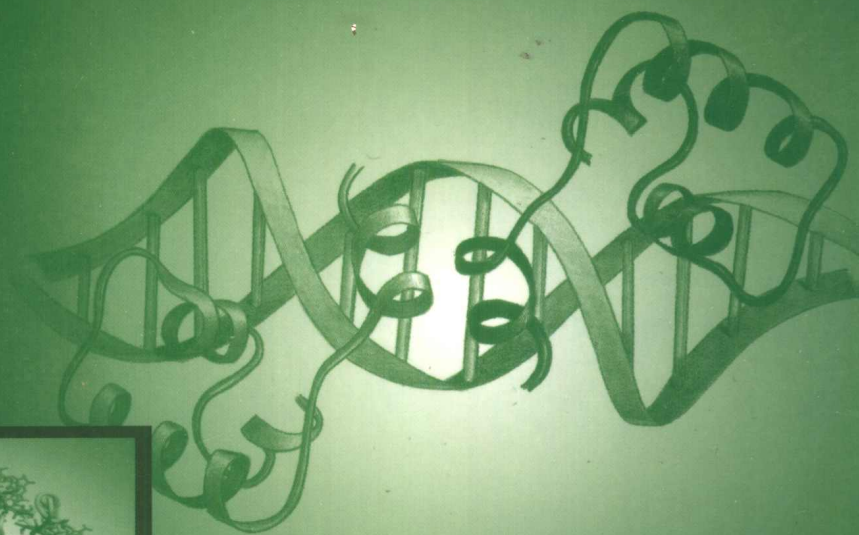


“十五”国家重点图书

现代生物技术丛书

蛋白质工程

王大成 主编



化学工业出版社

现代生物技术与医药科技出版中心

“十五”国家重点图书

现代生物技术丛书

7293
w31

蛋白质工程

王大成 主编

化学工业出版社
现代生物技术与医药科技出版中心
·北京·

(京) 新登字 039 号

图书在版编目(CIP)数据

蛋白质工程/王大成主编. —北京: 化学工业出版社,
2002.5
(现代生物技术丛书)
ISBN 7-5025-3837-2

I. 蛋… II. 王… III. 蛋白质-生物工程 IV. TQ93

中国版本图书馆 CIP 数据核字 (2002) 第 032990 号

现代生物技术丛书

蛋白质工程

王大成 主编

责任编辑: 叶 露 周 旭

责任校对: 陈 静

封面设计: 于 兵

*

化 学 工 业 出 版 社 出版发行
现代生物技术与医药科技出版中心

(北京市朝阳区惠新里 3 号 邮政编码 100029)

发行电话: (010) 64982530

<http://www.cip.com.cn>

*

新华书店北京发行所经销

北京市管庄永胜印刷厂印刷

三河市宇新装订厂装订

开本 787×1092 毫米 1/16 印张 17¼ 字数 410 千字

2002 年 9 月第 1 版 2002 年 9 月北京第 1 次印刷

ISBN 7-5025-3837-2/Q·21

定 价: 36.00 元

版权所有 违者必究

该书如有缺页、倒页、脱页者, 本社发行部负责退换

前 言

蛋白质工程是 20 世纪 80 年代初诞生的一个新兴生物技术领域，它一出现就以其在应用上的广阔前景和对分子生物学有关前沿研究的巨大推动而为世人瞩目，受到学术界和产业界的广泛重视。

生物技术的诞生和拓展，始终是与我们的生命现象本质认识的不断深入相伴而行的。现代分子生物学的杰出成就告诉我们，生命的主要物质基础是核酸和蛋白质，前者负责生命机体的世代遗传，后者是生命机体几乎所有重要活动的承担者。分子遗传学对核酸控制生命的精确了解，在 20 世纪 70 年代后期揭开了基因工程的序幕。进入 80 年代，结构生物学揭示了大量蛋白质分子的精确立体结构及其与复杂生物功能的关系，为设计改造天然蛋白质提供了蓝图；同时分子遗传学发展了以定位诱变为中心的基因操作技术，为通过基因修饰改造蛋白质提供了工具。这两方面的融合，促成了蛋白质工程这一新兴生物技术领域的诞生。1982 年，Winter 等首次报道通过基因定位诱变获得改性酪氨酸 tRNA 合成酶；1983 年，Ulmer 在“Science”上发表以“Protein Engineering”（蛋白质工程）为题的专论，一般将此视为蛋白质工程诞生的标志。

蛋白质工程的主要内容和基本目的可以概括为：以蛋白质分子的结构规律及其与生物功能的关系为基础，通过有控制的基因修饰和基因合成，对现有蛋白质加以定向改造，设计、构建并最终生产出性能比自然界存在的蛋白质更加优良、更加符合人类社会需要的新型蛋白质。

这是新一代的遗传工程。过去十多年，以重组 DNA 技术为中心的基因工程取得了重大进展，一批基因工程生产的药物制剂和工业用酶已相继进入商品市场，许多新的基因工程产品还在不断涌现，由此一个生物高技术产业蓬勃崛起。但是，基因工程原则上只生产自然界中已经存在的蛋白质。而大自然的创造远非完美无缺，无论对于社会生产还是对于人类生活，现存蛋白质都有许多不如人意之处。特别是作为基本的生命物质，天然蛋白质都是在生命机体中进化与发展的，经大自然精雕细刻所形成的优异结构与性能只是在这一特定环境中具有最适状态。而对人类的社会生产和生活，常常必须在非生命和非自然的状态下广泛应用蛋白质，这就必须对天然蛋白质加以适当的改造方可获得最佳的结构与性能。如许多工业用酶，必须改变天然酶的特性（如热稳定性、最适温度、最适 pH、对蛋白水解酶的抗性、底物结合特性、底物和反应的专一性等），方可适应生产和使用的需求。蛋白质工程研究通过定位的或有控制的基因修饰，提供改变蛋白质结构与性能的最有效的实用方法和技术途径，使天然蛋白质的改造成为现实可能。同时，结合蛋白质结构与功能关系的研究和蛋白质折叠机制的研究，通过基因合成，蛋白质工程还有可能制造结构和性能全新的蛋白质。以上述为基础，与基因工程有关技术的紧密结合，蛋白质工程正在形成以改造现有蛋白质和制造新型蛋白质为中心的第二代遗传工程。

与此同时，作为一种新型的强有力的研究手段，蛋白质工程对一些基本生物学问题的研究和解决，发挥了重要作用。从酪氨酸 tRNA 合成酶开始，现在已经广泛地运用蛋白质工程方法来研究各种蛋白质的结构与功能、蛋白质折叠、蛋白质分子设计等一系列基本分子生

物学问题。由于基因定位诱变的自如性，使这方面的工作突破了过去所用方法的局限，使有关研究达到前所未有的深度和广度。人和动植物几乎所有重要的功能性活动，诸如代谢、运动、发育、生长、思维、记忆、免疫防御、光合作用、固氮作用等都离不开蛋白质，所以蛋白质工程对揭示生命现象的本质和生命活动的规律具有重要意义，与人类的生产和生活密切相关。随着这一领域不断的开拓与发展，大量性能优良的新型蛋白质将被生产出来，并广泛应用于医药、工业、农业、环境保护等方面。

本书作为《现代生物技术丛书》中的一本，着重介绍与蛋白质工程有关的基础知识、主要方法和技术，以及一些具有典型意义的研究实例；力图使读者掌握蛋白质工程的基本原理，了解蛋白质工程的实际过程，熟悉从事蛋白质工程研究的重要方法。

全书共分六章。第一章介绍作为蛋白质工程蓝图的蛋白质结构原理，由王大成编写；第二章介绍蛋白质分子设计的方法，由徐筱杰编写；第三章的主要内容是进行蛋白质分子改造的分子生物学途径，由静国忠编写；第四章介绍突变蛋白质的物理化学性质分析，由曾宗浩编写；第五章概要介绍测定突变和天然蛋白质结构的主要方法，由李家瑶（X射线晶体结构分析）、王金凤（NMR溶液结构分析）、潘宪明（从氨基酸序列预测蛋白质结构）编写；第六章列举了一些蛋白质工程的应用实例，由静国忠编写。由于蛋白质工程是一个包含多种学科的边缘领域，因此本书各章节分别邀请了活跃在研究第一线的有关学科的专家执笔，感谢他们的共同努力使此书的编写得以顺利完成。

由于蛋白质工程的学科边缘性，至今对其应包含的内容并无明确的界定，加上编写者的水平有限，使本书在内容选裁和编写上也可能出现不当或错漏之处，敬望读者不吝指教。

感谢张英女士对本书部分图表制作和文字录入所给予的帮助。

王大成

2001年11月于北京

目 录

第一章 蛋白质结构基础	1
第一节 蛋白质结构的基本组件	1
一、20种天然常见氨基酸	1
1. 基本结构	1
2. 基本性质	2
3. 单一构型和旋转异构体	6
二、肽单位和多肽链	7
1. 肽链和肽单位	7
2. 多肽链的构象	8
三、 α 螺旋	10
1. 基本结构参数	10
2. 螺旋偶极子	11
3. 两亲性螺旋	12
4. 倾向于形成 α 螺旋的氨基酸	12
5. 3_{10} 螺旋和 π 螺旋	12
四、 β 层	13
1. β 链	13
2. 平行 β 层和反平行 β 层	13
五、环肽链	15
1. 回折	15
2. β 发夹	18
3. 环肽链的构象和优势氨基酸	20
六、疏水内核	20
第二节 蛋白质结构的组织和主要类型	21
一、蛋白质结构的层次体系	22
1. 一级结构	22
2. 二级结构	23
3. 结构模体	23
4. 结构域	26
5. 三级结构	28
6. 四级结构	28
二、蛋白质结构分类	29
1. 按结构域分类	29
2. 系统性分类	30
三、三类主要蛋白质结构	31

1. α 型结构	32
2. α/β 型结构	35
3. β 型结构	38
第三节 蛋白质结构的形成——多肽链的生物合成与折叠	42
一、多肽链的生物合成	43
1. 遗传密码与基因	43
2. 编码信息的转录	44
3. 编码信息的翻译	45
二、多肽链的折叠——蛋白质三维结构的形成	45
1. 蛋白质折叠的热力学基础	46
2. 蛋白质折叠的动力学因素	47
3. 帮助折叠的蛋白质和酶	49
三、蛋白质结构与蛋白质工程	51
1. 从预期结构到适配序列——蛋白质工程的基本困难	51
2. 改变蛋白质结构的一些途径	51
3. 增加蛋白质稳定性的基本途径	53
参考文献	55
第二章 蛋白质分子设计	56
第一节 基于天然蛋白质结构的分子设计	57
一、概述	57
二、蛋白质设计原理	59
三、蛋白质设计中的结构-功能关系研究	60
1. 根据结构信息确定残基的突变	61
2. 其他实验方法鉴定功能残基	62
3. 利用蛋白质同源性鉴定功能残基	63
四、天然蛋白质的剪裁	64
第二节 全新蛋白质设计	65
一、引言	65
二、蛋白质结构的从头设计	65
1. 二级结构模块单元的自组装	66
2. 配体诱导组装	72
3. 通过共价交叉连接实现肽的自组装	72
4. 在合成模板上肽的组装	73
5. 线性多肽折叠为球状结构	75
6. 基于组合库的全新蛋白质设计	77
三、蛋白质的功能设计	79
1. 通过反向拟合天然蛋白质设计新的功能	79
2. 键合及催化的从头设计	80
3. 在全新蛋白质中引入结合位点	80
4. 催化活性蛋白质的设计	82

5. 膜蛋白及离子通道的设计	82
6. 新材料的设计	84
第三节 计算蛋白质设计	84
一、能量表达	84
二、能量优化	85
三、序列优化	86
四、序列-结构专一性	87
五、底物专一性设计	87
六、金属结合位点的设计	87
参考文献	87
第三章 蛋白质的修饰和表达	88
第一节 蛋白质修饰的化学途径	88
一、功能基团的特异性修饰	88
1. 多位点取代	88
2. 单一的或限制性取代	89
3. 次级取代	91
二、基于蛋白质片段的嵌合修饰	92
1. 嵌合蛋白质——非共价缔合系统	92
2. 二硫键与嵌合蛋白质的形成	92
3. 嵌合蛋白质——通过化学激活形成肽键	93
4. 嵌合蛋白质——通过酶连接反应形成肽键	93
5. 通过非肽键形成嵌合蛋白质	94
第二节 蛋白质改造的分子生物学途径	94
一、编码基因的专一性位点和区域性定向突变	94
1. 编码基因的专一性位点突变	94
2. 区域性定向突变	101
二、基因融合和基因剪接	102
1. 利用基因融合技术表达外源基因的缘由	102
2. 基因融合的策略	102
3. 产生蛋白质分子嵌合体的方法	103
4. 蛋白质内含子介导的蛋白质分子间的连接	104
三、tRNA 介导定点掺入非天然氨基酸	105
第三节 重组蛋白质的表达	105
一、目标蛋白质在大肠杆菌中的表达	105
1. 表达载体的一般特点	106
2. 与外源基因有效表达的相关因素	108
3. 改善表达水平及溶解性的方法	111
二、目标蛋白质在酵母细胞中的表达	112
1. 有关酵母表达载体的复制、转录、筛选元件	112
2. 外源 mRNA 在酵母细胞中的翻译	114

3. 外源蛋白质在酵母中的分泌表达	114
4. 翻译后修饰	115
三、重组蛋白质在哺乳动物细胞中的表达	116
1. 选择哺乳动物细胞表达体系的优点	116
2. 两个主要的哺乳动物细胞表达系统	116
四、噬菌体显示	119
1. 关于表达载体	119
2. 噬菌体显示技术的操作	120
3. 噬菌体显示技术的应用	121
参考文献	121
第四章 突变蛋白质的物理化学性质分析	123
第一节 蛋白质溶液的热力学	123
一、热运动与蛋白质构象	124
二、热力学函数与热力学平衡	126
三、热容量	128
四、van't Hoff 焓	129
五、折叠/退折叠转变	129
六、量热法与折叠过程热力学	131
第二节 蛋白质折叠动力学	133
一、折叠动力学研究技术	134
二、两态动力学	136
三、过渡态	138
1. 折叠过程与过渡态	138
2. 对折叠过渡态的性质分析	139
四、折叠中间态	142
1. 熔球态	142
2. 快态与慢态	143
3. 二硫键引起的中间态	144
4. 多结构域蛋白的折叠	145
五、折叠的基本过程	147
1. 接触形成	147
2. 螺旋-链环转变	148
3. β 发卡形成	148
第三节 突变、稳定性和折叠	149
一、热力学参数在分子水平上的解释	150
1. 静电相互作用	150
2. 范德华相互作用	151
3. 氢键	152
4. 疏水效应	152
5. 二硫键	154

二、突变与热稳定性	155
1. 疏水突变	156
2. 氢键突变	157
3. 适应极端条件的突变体	159
三、突变与折叠过程	161
1. 过渡态的突变分析	161
2. 突变对稳定性和折叠过程的影响	164
参考文献	165
第五章 天然和重组蛋白质结构测定	166
第一节 X射线晶体结构分析	166
一、概念——基本原理	166
1. X射线	166
2. X射线衍射	167
3. 周期函数和富里哀定理	168
4. 晶体	169
5. 天然和重组蛋白质X射线晶体结构分析的特殊性	169
二、蛋白质结晶和晶体生长	171
1. 对原材料的要求	172
2. 晶体生长的生物化学条件	173
3. 晶体生长的物理条件	173
4. 技术和方法	174
5. 蛋白质晶体的鉴定	174
三、衍射数据收集	175
1. 晶体的收集和处理	175
2. X射线源的选择	175
3. 衍射线记录装置及其使用方法	176
4. 衍射数据收集的全自动化考虑	177
四、确定位相	178
1. 数据的统一和还原	178
2. 同晶置换法	179
3. 分子置换法	182
4. 差值电子密度法	183
5. 多波长反常散射法	183
6. 全自动化数据处理和位相计算	185
五、电子密度图诠释	185
1. 电子密度修饰	185
2. 分辨率	186
3. 分析电子密度图	186
4. 全自动化程序发展前景	188
六、结构模型精化	188

1. R 因子	188
2. 限制性最小二乘修正	189
3. 全自动化软件	189
4. 结构模型表达	189
参考文献	190
第二节 核磁共振波谱的溶液结构解析	190
一、概述——蛋白质溶液三级结构测定	191
1. 蛋白质的结构信息	191
2. 核磁共振的波谱信息	192
3. 由波谱参数提取蛋白质的结构信息	193
4. 由核磁共振数据计算蛋白质的三级结构	196
二、基本的多维核磁共振实验	198
1. 二维同核 (^1H - ^1H) 核磁共振实验	199
2. 二维异核 (^1H - ^{15}N / ^{13}C) 核磁共振实验	203
3. 三维异核 (^1H - ^{13}C - ^{15}N) 核磁共振实验	205
4. 四维异核 (^1H - ^{13}C - ^{15}N) 核磁共振实验	211
5. TROSY 类型的多维核磁共振实验	213
三、多维核磁共振波谱解析简介	215
1. 核磁共振谱峰的序列指认	215
2. 立体定向 (stereospecific) 指认	217
3. NOE 交叉峰的指认及峰强度的正确测定	220
参考文献	224
第三节 蛋白质结构预测	225
一、蛋白质结构预测方法准确性的评估	227
二、序列相似性对比	227
三、蛋白质二级结构预测	231
1. 统计方法	232
2. 基于已有知识的预测方法 (knowledge-based method)	233
3. 混合方法 (hybrid system method)	233
四、蛋白质三级结构预测	234
五、蛋白质结构预测与药物设计	236
六、蛋白质数据库	237
1. 核酸序列数据库	237
2. 蛋白质序列数据库	237
3. 蛋白质结构数据库 (PDB)	237
4. 蛋白质二级结构数据库 (DSSP)	239
参考文献	239
第六章 蛋白质工程在医药工业中的应用	241
第一节 医用抗体的蛋白质工程	242
一、抗体的产生和抗体的结构简介	243

1. 多肽构件的随机组合产生大量不同抗体	243
2. 抗原结合的特异性是由超可变结构域决定的	243
3. 恒定区的作用是稳定抗体分子	244
4. 稳定区介导效应器功能	244
二、从抗血清到重组抗体	244
三、抗体工程	245
1. 鼠-人嵌合抗体	245
2. 来自鼠抗体可变结构域的骨架区可被人源化	246
3. 抗体的三维结构可在计算机上建模	247
第二节 组织纤维蛋白溶酶原激活因子的蛋白质工程	248
一、有关重组 t-PA	248
二、产生 t-PA 突变体的基本原理	248
三、t-PA 蛋白质工程的几个方面	249
1. 减慢清除的 t-PA 变体	249
2. 血纤维特异性	250
第三节 基于蛋白质结构的小分子药物设计	251
一、用 DOCK 发现先导化合物	252
二、Haloperidol 衍生物对 HIV-1 PR 的抑制	252
参考文献	253
中西文名词对照	254

第一章 蛋白质结构基础

蛋白质工程的基本目标是，按预期的结构和功能，通过基因修饰或基因合成，对现有蛋白质加以定向改造，设计、构建并最终生产出性能比天然蛋白质更加优良、更加符合人类社会需要的新型蛋白质。无论是改造现有蛋白质还是从头设计全新蛋白质，都必须以蛋白质分子的结构规律及其与生物功能的关系为基础，以天然蛋白质分子的三维结构为基本蓝图。20世纪分子生物学的基本成就已使人们普遍认识到，任何一个蛋白质，在其自然状态或活性形式下，都具有特征而稳定的三维结构（空间结构，立体结构）；一旦这种专一的空间结构遭到破坏，即使化学结构完全不变，蛋白质的功能也会立即消失。而且，在执行正常功能时，这种专一的三维结构常常必须发生微妙的运动。没有特征的三维结构就没有复杂的蛋白质功能，具有独特的三维结构是功能蛋白质分子最基本的属性。因此，以重新设计蛋白质分子为目标的蛋白质工程研究者必须深刻了解和把握蛋白质结构特别是三维结构知识。

根据2001年10月国际蛋白质数据库公布的资料，在原子或接近原子分辨率水平解析的蛋白质三维结构数量已超过16000。大量精细结构的阐明显示，蛋白质三维结构具有极大的多样性和极高的复杂性。与DNA分子大体都具有双螺旋式结构不同，不同类型的蛋白质具有完全不同的分子形貌，真是“千人千面”，它们各具特征的结构组织与丰富多彩的功能机理密切联系，构成蛋白质功能多样性的结构基础。与此同时，大量蛋白质结构的测定也揭示不同蛋白质分子结构之间存在的共同特征，包括基本的结构组件、多层次的结构组织方式、在结构域水平上的结构类型等。本章将概要介绍这些具有共性的蛋白质结构知识，作为蛋白质工程的共同基础。

第一节 蛋白质结构的基本组件

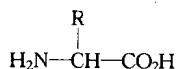
尽管蛋白质具有极其复杂的总体结构，但在化学上它们都是由20种天然氨基酸按特定的顺序通过肽键连接形成的具有有限长度的多肽链。不同蛋白质间最基本的差别就是其组成多肽链的氨基酸序列和长度不同。尽管以此为基础形成的蛋白质的总体三维结构极其复杂和多样，但对大量蛋白质高分辨率精细三维结构的分析显示，在各类蛋白质结构中存在一些共同的组件，它们构成蛋白质分子组织的基本要素，也是设计和构建新蛋白质的主要基础。本节将对目前已知的蛋白质结构的化学的和空间的基本组件作一概要介绍。

一、20种天然常见氨基酸

构成蛋白质分子的基本化学单位是氨基酸。在所有蛋白质中常见的天然氨基酸共20种，在有机体内通过生物合成连接成多肽链，其顺序由编码基因中的核苷酸三联体遗传密码决定（参见第三节）。

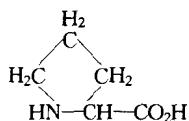
1. 基本结构

20种常见氨基酸中，19种都具有如下共同的化学结构：



其中心碳原子（C^α）除结合一个H原子外，还分别连接一个氨基（NH₂）和一个羧基

(COOH)，构成氨基酸的主链，它们在所有氨基酸中都是相同的。R 称为氨基酸的侧链，不同氨基酸的区别就是其侧链 R 的化学结构不同。另有一种脯氨酸具有类似而不相同的化学结构，它们的侧链与主链 N 原子共价结合，形成一个亚氨基酸：



20 种常见氨基酸的侧链结构如图 1-1 所示。上列结构式中的中心碳原子称为 α 碳 (C^α)，侧链原子按离 C^α 原子距离顺序称为 β 、 γ 、 δ 、 ϵ 和 ξ ，但通常侧链中的基团是按与它键合的 C 原子的名称来定名的，如赖氨酸中的 ϵ 氨基 (见图 1-1)。在实用中，这些氨基酸有通用的三字符和单字符缩写，见图 1-1 和表 1-1 所示，它们也将在本书中使用。当这些氨基酸连接成多肽链时，其主链、侧链共同作为一个基本单位称为氨基酸残基 (amino acid residue)。本书在涉及蛋白质结构的所有地方，“氨基酸”均指“氨基酸残基”，有时为文字简洁而略去“残基”二字。

除脯氨酸外，甘氨酸(Gly)的结构也有特殊性，它的侧链只有一个氢原子(见图 1-1)。缺乏长侧链的制约使多肽链在 Gly 残基出现的地方具有显著的构象柔性 (flexibility)。在 Gly 中 C^α 的非对称性消失，使其没有 L 型和 D 型构型之分，可以在广泛的构象空间中出现。

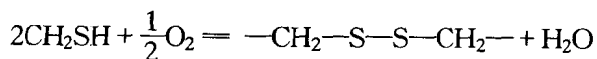
2. 基本性质

20 种不同的氨基酸侧链具有各种各样的化学特性，当它们以各种顺序结合成蛋白质分子时，其性质会有更多的变化，使其蛋白质的特性远不同于简单的有机分子。因此，生物活性蛋白质中的氨基酸残基的化学、物理性质常常会与游离氨基酸有显著不同。尽管蛋白质的化学性质并不等同于其组成氨基酸的加合，但 20 种氨基酸侧链的性质是了解蛋白质的基本出发点。

20 种氨基酸的一些基本化学性质和物理性质概列于表 1-1 和表 1-2 中。以下简要地分类介绍这些氨基酸的一些性质。

(1) 疏水氨基酸 疏水氨基酸包括 Ala、Val、Leu、Ile、Met、Pro、Phe。这类残基的侧链一般都没有化学反应性，其共同的特性是疏于与水相互作用，趋于彼此间或与其他非极性原子相互作用。所有蛋白质分子都有一部分这类残基密堆积在内部，形成疏水内核，这是稳定蛋白质三维结构的主要因素。这类残基的疏水相互作用被认为是多肽链折叠 (folding) 的原初推动力。其中 Pro 残基并不常出现在分子内部，它的侧链与主链 N 原子共价相联，形成一个五元环，对主链 N- C^α 的旋转形成制约，并使其与前一残基形成的肽键易于成顺式 (cis)。因此这一残基可对各肽链的构象发生重要影响。一般 Pro 残基五元环的 C^α 、 C^β 、 C^δ 和 N 处于一近似的平面中，而 C^γ 偏离平面大约 0.5° 。

(2) 极性氨基酸 极性氨基酸包括 Ser、Thr、Asn、Gln、Cys、His、Tyr 和 Trp。它们的侧链都含有极性基团，可以是氢键的给体或受体，并且有不同程度的化学反应性。在蛋白质三维结构中靠近的两个 Cys 常常可被氧化形成二硫键，作为一个合成的结构单位成为胱氨酸 (Cystine) (图 1-2)。通常二硫键按以下反应式由空气氧化产生最终产物：



这一反应需要一个氧化环境，因此这种二硫键一般并不存在于细胞内蛋白质中，因为那里主要是一个还原环境。二硫键常常出现在由细胞分泌出来的胞外蛋白质中，在真核生物中二硫

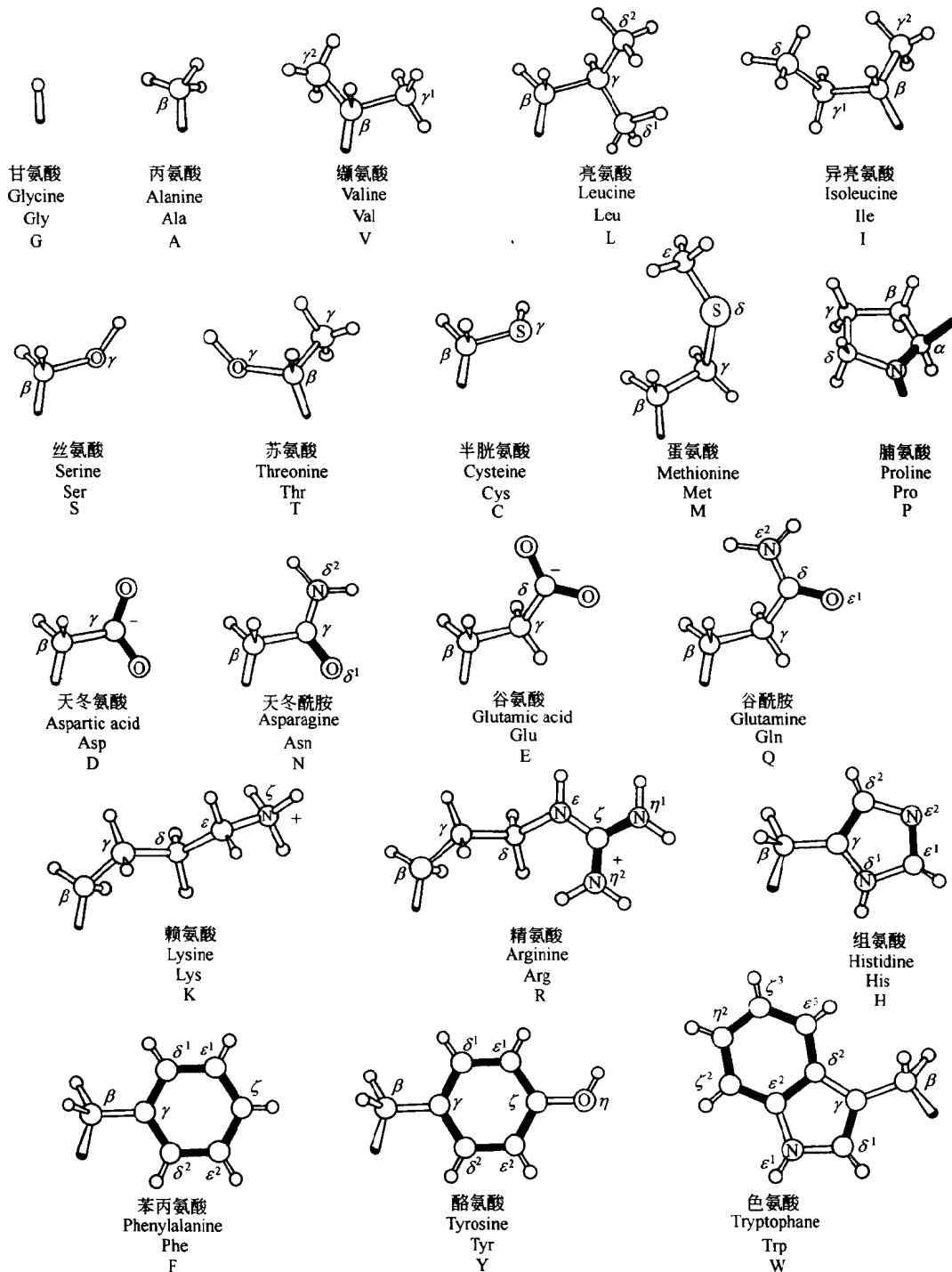


图 1-1 天然蛋白质中常见 20 种氨基酸的侧链

图中未标符号的小球和大球分别是氢原子和碳原子，双键和局部双键分别以黑色和阴影示之，在脯氨酸中粗黑键表示属于主链的部分。在氨基酸全名之下，是常用的三字符和单字符缩写
 对异亮氨酸和苏氨酸，图中示出的只是在天然蛋白质中存在的单一异构体

表 1-1 氨基酸残基的一些立体化学相关性质

残基	质量/Da ^①	Van de Waals 体积 ^② /nm ³	可及表面积 ^③ /nm ²			
			总体	主链原子	侧链总体	侧链极性原子
丙氨酸 Ala(A)	71.09	0.067	1.13	0.46	0.67	
精氨酸 Arg(R)	156.19	0.148	2.41	0.45	1.96	1.07
天冬酰胺 Asn(N)	114.11	0.096	1.58	0.45	1.13	0.69
天冬氨酸 Asp(D)	115.09	0.091	1.51	0.45	1.06	0.58
半胱氨酸 Cys(C)	103.15	0.086	1.40	0.36	1.04	0.69
谷酰胺 Gln(Q)	128.14	0.114	1.89	0.45	1.44	0.91
谷氨酸 Glu(E)	129.12	0.109	1.83	0.45	1.38	0.77
甘氨酸 Gly(G)	57.05	0.048	0.85	0.85		
组氨酸 His(H)	137.14	0.118	1.94	0.43	1.51	0.49
异亮氨酸 Ile(I)	113.16	0.124	1.82	0.42	1.40	
亮氨酸 Leu(L)	113.16	0.124	1.80	0.43	1.37	
赖氨酸 Lys(K)	128.17	0.135	2.11	0.44	1.67	0.48
蛋氨酸 Met(M)	131.19	0.124	2.04	0.44	1.60	0.43
苯丙氨酸 Phe(F)	147.18	0.135	2.18	0.43	1.75	
脯氨酸 Pro(P)	97.12	0.090	1.43	0.38	1.05	
丝氨酸 Ser(S)	87.08	0.073	1.22	0.42	0.80	0.36
苏氨酸 Thr(T)	110.11	0.093	1.46	0.44	1.02	0.28
色氨酸 Trp(W)	186.21	0.163	2.59	0.42	2.17	0.27
酪氨酸 Tyr(Y)	163.18	0.141	2.29	0.42	1.87	0.43
缬氨酸 Val(V)	99.14	0.105	1.60	0.43	1.17	
平均值	119.40	0.161				

① 1Da (道尔顿) = 1u (原子质量单位) = $(1.660\ 540\ 2 \pm 0.000\ 001\ 0) \times 10^{-27}$ kg。

② Richards F M. J Mol Biol, 1974, 82: 1~14。

③ Miller S, et al. J Mol Biol, 1987, 196: 641~656。

表 1-2 氨基酸残基的一些物理化学相关性质

残基	可解离集团 ^① pK 值	亲水性 ^② /kJ·mol ⁻¹	疏水性/kJ·mol ⁻¹		在蛋白质中 出现频率 ^④ /%
			侧链类似物 ^③	氨基酸 ^③	
精氨酸 Arg	12.0	-97.44	66.29	12.54	5.7
天冬氨酸 Asp	3.9~4.0	-55.76	40.38	10.45	5.3
谷氨酸 Glu	4.3~4.5	-52.79	32.40	10.45	6.2
天冬酰胺 Asn		-50.45	31.68	0.84	4.4
赖氨酸 Lys	10.4~11.1	-49.78	27.13	12.54	5.7
谷酰胺 Gln		-49.20	27.09	0.84	4.0
组氨酸 His	6.0~7.0	-52.92	23.41	2.09	2.2
丝氨酸 Ser		-31.14	18.14	1.25	6.9
苏氨酸 Thr		-30.39	14.67	1.67	5.8
酪氨酸 Tyr	10.0~10.3	-35.53	4.51	9.61	3.2
甘氨酸 Gly		0	0	0	7.2
脯氨酸 Pro		—	—	-5.85	5.1
半胱氨酸 Cys	9.0~9.5	-15.17	-1.42	-4.18	1.7
丙氨酸 Ala		-1.88	-3.64	-2.09	8.3
色氨酸 Trp		-34.57	-5.81	-14.21	1.3
蛋氨酸 Met		-16.18	-5.89	-5.43	2.4
苯丙氨酸 Phe		-13.17	-8.53	-10.45	3.9
缬氨酸 Val		-16.72	-12.96	-6.27	6.6

残基	可解离集团 ^① pK 值	亲水性 ^② /kJ·mol ⁻¹	疏水性/kJ·mol ⁻¹		在蛋白质中 出现频率 ^④ /%
			侧链类似物 ^②	氨基酸 ^③	
异亮氨酸 Ile		-1.00	-16.64	-7.52	5.2
亮氨酸 Leu		-0.46	-16.64	-7.52	9.0
α -氨基	6.8~8.0				
α -羧基	3.5~4.3				

① Matthews J B, et al. CRC Crit Rev Biochem, 1985, 18: 91~197。

② Radzicka A, Wolfenden R. Biochemistry, 1988, 27: 1664~1670。

③ Levitt M. J Mol Biol, 1976, 104: 59~107。

④ McCalon P, Argos P. Proteins, 1988, 4: 99~122。

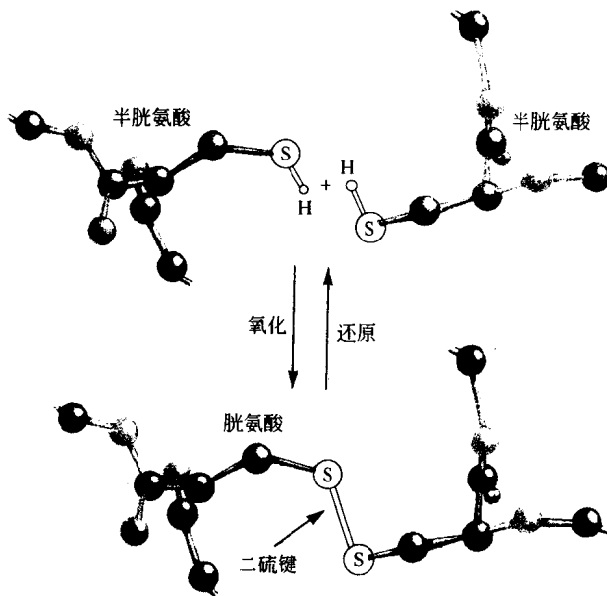


图 1-2 从两个半胱氨酸残基形成二硫键

键式在内质网的间隙中形成。二硫键可以稳定蛋白质的三维结构。在一些蛋白质分子中它可以不同的多肽链连接在一起(如胰岛素中的 A 链、B 链);更经常出现在分子内的二硫键可稳定单肽链的折叠,使蛋白质不易被降解。特别对许多小蛋白质分子,如蝎毒素、蛋白酶抑制剂,常常需要多对二硫键作为重要的稳化因素使其三维结构得以稳定。在蛋白质工程中已有很多的工作试图通过定位突变在酶分子中引入二硫键,以提高它们的热稳定性,应用于工业生产。

Tyr 和 Trp 连同非极性残基 Phe 都具有芳香性侧链,是使蛋白质产生紫外线吸收和荧光特性的主要因素。这些残基的光谱特性及紫外线吸收特性分别示于表 1-3 和图 1-3。它们对于介质环境非常敏感,常常作为蛋白质结构变化的非常有用的探针。

表 1-3 氨基酸残基的一些物理化学相关性质^①

氨基酸	光吸收		荧光发射	
	λ_{max}/nm	摩尔吸收/ $mol^{-1}\cdot cm^{-1}$	λ_{max}/nm	量子产率
苯丙氨酸	257.4	197	282	0.04
酪氨酸	274.6	1420	303	0.21
色氨酸	279.8	5600	348	0.20

① Teale F W J, Weber H. Biochem J, 1957, 65: 476~482。