



中央民族大学国家“十五”“211工程”建设项目



DANBAIZHI GONGCHENG YUANLI JI YINGYONG

黄耀江 王 琰 冯健男 编著

蛋白质工程原理及应用

中央民族大学出版社



中央民族大学国家“十五”“211工程”建设项目

黄耀江 王 琰 冯健男 编著

蛋白质工程原理及应用

中央民族大学出版社

图书在版编目 (CIP) 数据

蛋白质工程原理及应用/黄耀江等编著. —北京:
中央民族大学出版社, 2007.8

ISBN 978-7-81108-445-0

I. 蛋… II. 黄… III. 蛋白质—生物工程 IV. TQ93

中国版本图书馆 CIP 数据核字 (2007) 第 138142 号

蛋白质工程原理及应用

编 著 黄耀江 王 琰 冯健男

责任编辑 杨 玉

美术编辑 布拉格

出 版 者 中央民族大学出版社

北京市海淀区中关村南大街 27 号 邮编:100081

电话:68472815(发行部) 传真:68932751(发行部)

68932218(总编室) 68932447(办公室)

发 行 者 全国各地新华书店

印 刷 者 北京宏伟双华印刷有限公司

开 本 880×1230(毫米) 1/32 印张:15.875

字 数 398 千字

印 数 2000 册

版 次 2007 年 8 月第 1 版 2007 年 8 月第 1 次印刷

书 号 ISBN 978-7-81108-445-0

定 价 32.00 元

版权所有 翻印必究

前 言

生物工程亦称生物技术,主要包括基因工程、蛋白质工程(包括酶工程)、细胞工程、微生物工程(即发酵工程)等技术体系。作为 20 世纪 70 年代初发展起来的一个新兴技术领域,生物技术已成为世界新技术革命的重要组成部分,在不断提高人类的生活质量、改善人类赖以生存的环境、改善人类健康状况以及治疗严重疾病等方面,生物工程起到至关重要的作用。

作为生物技术的重要组成部分,蛋白质工程是在重组 DNA 技术应用于蛋白质结构与功能研究之后发展起来的一门新兴学科。正当基因工程诞生 10 周年之际,著名科学家 Kevin M. Ulmer(额尔默)于 1983 年在 SCIENCE 发表了一篇文章《Protein Engineering》,该论文的发表是蛋白质工程诞生的重要标志。所谓蛋白质工程,就是通过对蛋白质已知结构和功能的了解,借助计算机辅助设计,利用基因定位诱变等技术,特异性地对蛋白质结构基因进行改造,产生具有新的特性的蛋白质的技术,并由此深入研究蛋白质的结构与功能的关系,使蛋白质更好地造福于人类。由于蛋白质工程开创了按照人类意愿设计制造符合人类需要的蛋白质的新时期,因此被誉为第二代遗传工程。蛋白质工程的出现,为认识和改造蛋白质分子提供了强有力的手段。

本书是作者在蛋白质工程领域多年的科研与教学工作基础之上、参考国内外最新出版的著作及文献资料编写而成的。主要内容为蛋白质结构、蛋白质组学、蛋白质生物信息学、蛋白质的分子模拟与药物设计、蛋白质表达、蛋白质分离纯化工程,以及蛋白质工程的应用如抗体工程、蛋白多肽药物、酶蛋白工程、蛋白

芯片等。本书既可作为生物工程、生态环境、医学、药学等专业高年级本科生、研究生的教材，亦可供从事生物工程研究和生产的技术人员参考。

全书由黄耀江主编，负责结构设计、章节安排、内容调整及全书审核等工作。本书共分为十一章，黄耀江编写第一、第二、第四章及第五、第六、第八章的部分内容；军事医学科学院冯健男编写第三章；沈阳农业大学张少斌编写第五、第十一章；南京农业大学的瞿桂香编写第六章；海军总医院王琰编写第七章；北京军区总医院金华编写第八章；中央民族大学刘丽娜、米佳编写第九章；沈阳农业大学阚玉国编写第十章。本书得到了中央民族大学国家“十五”、211 工程建设项目经费的支持。在全书编写过程中，作者参考了大量国内外相关的书籍和文献，在此向这些前辈和同行们表示衷心的感谢！

由于生物技术的发展迅速和编写时间十分紧张，许多蛋白质工程的新技术、新成果来不及编入本书，加之作者水平有限，错误及不妥之处在所难免，恳请前辈、同行以及读者提出宝贵意见。

作者

2007 年暑期于中央民族大学

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
第二节 蛋白质结构的组成和主要类型	12
第三节 蛋白质结构的形成——多肽链的链的生物合成 与折叠	22
第四节 蛋白质的结构与功能关系	36
第二章 蛋白质组学、蛋白质的生物信息学	43
第一节 蛋白质组学原理与技术	43
第二节 蛋白质的生物信息学资源	52
第三节 蛋白质的生物信息学分析	58
第三章 蛋白质结构模拟与药物设计	69
第一节 计算机辅助蛋白质分子设计	69
第二节 蛋白质结构数据库	70
第三节 蛋白质结构模拟	78
第四节 蛋白质分子设计	85
第五节 药物分子设计	93
第四章 蛋白质的分子改造、修饰和表达	107
第一节 蛋白质修饰的化学途径	107
第二节 蛋白质改造的分子生物学途径	115
第三节 重组蛋白质的表达	130
第五章 蛋白质的分离、纯化和鉴定	146

第一节	蛋白质分离纯化概述	147
第二节	蛋白质的提取	149
第三节	蛋白质的粗分级	154
第四节	蛋白质的细分级	165
第五节	蛋白质的鉴定	199
第六章	天然和重组蛋白质的结构测定	207
第一节	X 射线晶体结构分析	207
第二节	核磁共振波谱的溶液结构解析	219
第三节	蛋白质的结构预测	231
第七章	抗体工程	244
第一节	概述	244
第二节	抗体分子的结构与功能	247
第三节	细胞工程抗体	262
第四节	基因工程抗体	270
第八章	蛋白质芯片	299
第一节	蛋白芯片的概念和原理	299
第二节	蛋白芯片分类和意义	303
第三节	蛋白芯片制备和检测	305
第四节	蛋白芯片技术的应用	316
第五节	蛋白芯片存在的问题及技术展望	331
第九章	蛋白质和多肽药物工程	353
第一节	概述	353
第二节	促红细胞生成素	354
第三节	组织型纤溶酶原激活剂	355
第四节	白细胞介素	358
第十章	酶蛋白工程	362
第一节	酶学的基础知识	362
第二节	酶工程的基础知识	383

第三节	微生物酶的发酵生产	391
第四节	酶的提取、分离和纯化	412
第五节	酶分子的修饰	421
第六节	酶和细胞的固定化	429
第七节	酶的反应器和酶的传感器	445
第八节	有机介质中的酶促反应	456
第九节	酶制剂的应用	459
第十一章	蛋白质工程在农牧业的应用	480
第一节	植物抗虫蛋白质工程	480
第二节	植物抗除草剂蛋白质工程	487
第三节	植物品种改良蛋白质工程	489
第四节	植物抗逆、抗病蛋白质工程	492
第五节	与农牧业相关的其他蛋白质工程	494

1. 蛋白质含量的测定：蛋白质含量测定方法很多，如凯氏定氮法、双缩脲法、Lowry 法等。其中凯氏定氮法是最常用的方法，其原理是测定样品中的总氮含量，然后根据氮含量计算出蛋白质含量。

第一章 蛋白质结构基础

(氨基酸) $25.0 \times$ 质量含氮测品样 = 质量蛋白质

生物体是由多种复杂成分组成的。其中有许多的有机分子和金属、非金属离子，也有各种各样的生物大分子，如多糖、脂质、蛋白质和核酸等。这些生物大分子与一般的合成大分子最大的区别在于特定结构层次和时空特性。基于生物体内蛋白质种类的极其多样性、分布广泛性以及所担负任务的重要性，本章将对蛋白质的结构基础进行阐述。

蛋白质是一种含有由 DNA 编码的 20 种 L 型 α 氨基酸，通过 α 碳原子上的取代基间形成的酰胺键连成的，具有特定空间构象和生物功能的肽链构成的生物大分子。只含有肽链的蛋白质是简单蛋白，肽链和其他的组分还能形成复合蛋白。应该将蛋白质和肽链两者区分开，蛋白质和对应的肽链的差别在于折叠的方式。一条肽链只有通过折叠成特定的空间构象后，才能称为蛋白质。因此蛋白质是经折叠后具有特定空间构象的肽链；肽链是去折叠 (unfolding)、无特定空间构象的蛋白质。常见的松散肽链，或是从核糖体上释放出来的新生肽链，或是蛋白质经较剧烈条件处理后得到的肽链，后者习惯上也称为变性蛋白。

第一节 蛋白质结构的基本组件

一、蛋白质的元素组成

经元素分析已知，单纯蛋白质的元素组成为：碳 50% ~ 55%，氢 6% ~ 7%，氧 19% ~ 24%，氮 13% ~ 19%，大多数还含有 4% 以下

的硫。自然界的蛋白质其含氮量比较接近,平均含氮量约为 16%。由于测定生物样品中的含氮量比测定其中蛋白质的量容易很多,因此,可以通过测定样品中氮的含量计算其中蛋白质含量。kjedahl 定氮法正是依此来计算蛋白质含量的,公式如下:

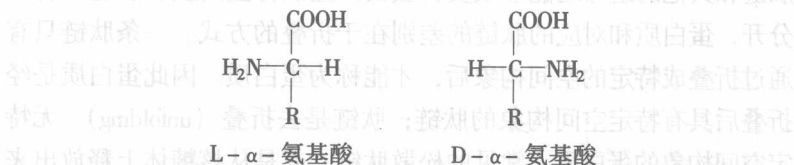
$$\text{蛋白质含量} = \text{样品所测含氮量} \times 6.25 \text{ (蛋白系数)}$$

二、蛋白质的基本结构单位——氨基酸 (Amino acid)

蛋白质种类虽然繁多,但其经酸、碱或蛋白酶催化作用后,最后得到的水解产物都是氨基酸。因此,氨基酸是蛋白质的基本结构单位。

(一) 氨基酸的结构

自然界中的氨基酸有 300 余种,但组成人体蛋白质的基本氨基酸仅有 20 种,这些氨基酸的结构各不相同,但结构中的氨基 ($-\text{NH}_2$) 都与邻接羧基 ($-\text{COOH}$) 的 α -碳原子相连接,故它们都属于 α -氨基酸,结构如下:



R 为侧链。其特点是:除 R 为氢原子 (甘氨酸) 外,其他氨基酸的 α -碳原子都是不对称碳原子 (手性碳原子),因而可形成不同构型,具旋光性质;除脯氨酸为亚氨基酸外,其余 19 种均符合上述通式;不同氨基酸的 R 侧链各异,它们的分子量、解离程度和化学反应性质也不相同。

(二) 氨基酸的分类

前面说过,组成人体蛋白质的 20 种基本氨基酸区别就在于侧链 R 基的不同,因此,按 R 基的化学结构和理化性质的不同主要分为四类 (见表 1-1):

表 1-1 氨基酸分类

极性 状况	带电荷 状况	化学结构式	氨基酸名 称及简写	PK ₁ (α-COOH)	PK ₂ (α-NH ₂)	PK(R)	等电点 (PI)
极性 氨基酸	酸性	$\text{HO}-\text{CH}_2-\text{CH}-\text{COO}^-$ NH_3^+	丝氨酸, Ser, S	2.21	9.15		5.68
	碱性	$\text{CH}_3-\text{CH}-\text{CH}-\text{COO}^-$ OH NH_3^+	苏氨酸, Thr, T	2.63	10.43		6.53
	中性	$\text{H}_2\text{N}-\text{C}(=\text{O})-\text{CH}_2-\text{CH}-\text{COO}^-$ NH_3^+	天冬酰胺, Asn, N	2.02	8.8		5.41
	中性	$\text{H}_2\text{N}-\text{C}(=\text{O})-\text{CH}_2-\text{CH}_2-\text{CH}-\text{COO}^-$ NH_3^+	谷氨酰胺, Gln, Q	2.17	9.13		5.65
	酸性	$\text{HO}-\text{C}_6\text{H}_4-\text{CH}_2-\text{CH}-\text{COO}^-$ NH_3^+	酪氨酸, Tyr, Y	2.20	9.11	10.07 (OH)	5.66
非极性 氨基酸	酸性	$\text{HS}-\text{CH}_2-\text{CH}-\text{H}_3$ N	半胱氨酸, Cys, C	1.71	10.78	8.33 (SH)	5.02

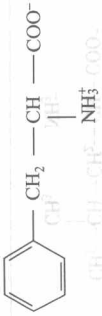
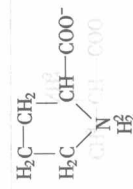
续表

极性状况	带电荷状况	化学结构式	氨基酸名称及简写	PK ₁ (α-COOH)	PK ₂ (α-NH ₂)	PK(R)	等电点 (PI)
酸性氨基酸	酸性氨基酸带负电荷	$\text{OOC}-\text{CH}_2-\underset{\text{HN}_3^+}{\text{CH}}-\text{COO}^-$	天冬氨酸, Asp, D	2.09	9.82	3.86 (βCOOH)	2.97
		$\text{OOC}-\text{CH}_2-\underset{\text{HN}_3^+}{\text{CH}}-\text{COO}^-$	谷氨酸, Glu, E	2.19	9.67	4.25 (γCOOH)	3.22
	碱性氨基酸带正电荷	$\text{H}^+\text{N}^+ \text{C}(\text{NH}_3^+)=\text{CH}_2-\text{CH}-\text{COO}^-$	组氨酸, His, H	1.82	9.17	6.00 (咪唑基)	7.59
中性氨基酸	中性氨基酸带正电荷	$\text{H}_3\text{N}^+-\text{CH}_2-\text{CH}_2-\text{CH}_2-\underset{\text{NH}_3^+}{\text{CH}}-\text{COO}^-$	赖氨酸, Lys, K	2.18	8.95	10.53 (εNH ₃)	9.74
		$\text{H}_2\text{N}-\underset{\text{NH}_3^+}{\text{C}}-\text{NH}-\text{CH}_2-\text{CH}_2-\text{CH}_2-\text{CH}-\text{COO}^-$	精氨酸, Arg, R	2.17	9.04	12.48 (胍基)	10.76

续表

极性 状况	带电荷 状况	化学结构式	氨基酸名 称及简写	PK ₁ (a - COOH)	PK ₂ (a - NH ₂)	PK(R)	等电点 (PI)
非极性氨基酸		$\begin{array}{c} \text{H} - \text{CH} - \text{COO}^- \\ \\ \text{NH}_3^+ \end{array}$	甘氨酸, Gly, G	2.34	9.60		5.97
		$\begin{array}{c} \text{CH}_3 - \text{CH} - \text{COO}^- \\ \\ \text{NH}_3^+ \end{array}$	丙氨酸, Ala, A	2.34	9.69		6.02
		$\begin{array}{c} \text{CH}_3 - \text{CH} - \text{CH} - \text{COO}^- \\ \quad \\ \text{CH}_3 \quad \text{NH}_3^+ \end{array}$	缬氨酸, Val, V	2.32	9.62		5.97
		$\begin{array}{c} \text{CH}_3 - \text{CH} - \text{CH}_2 - \text{CH} - \text{COO}^- \\ \quad \quad \\ \text{CH}_3 \quad \quad \text{NH}_3^+ \end{array}$	亮氨酸, Leu, L	2.36	9.60		5.98
		$\begin{array}{c} \text{CH}_3 - \text{CH}_2 - \text{CH} - \text{CH} - \text{COO}^- \\ \quad \quad \\ \text{CH}_3 \quad \quad \text{NH}_3^+ \end{array}$	异亮氨酸, Ile, I	2.36	9.68		6.02

续表

极性 状况	带电荷 状况	化学结构式	氨基酸名 称及简写	PK ₁ (a-COOH)	PK ₂ (a-NH ₂)	PK(R)	等电点 (PI)
非 极 性 氨 基 酸			苯丙氨酸, Phe, F	1.83	9.13		5.48
			甲硫氨酸 (蛋氨酸), Met, M	2.28	9.21		5.75
			脯氨酸, Pro, P	1.99	10.60		6.30
弱 极 性 氨 基 酸			色氨酸, Trp, W	2.38	9.38		5.89

1. 非极性疏水氨基酸

其特征是含有非极性 R 侧链, 它们显示出不同程度的疏水性。这一组共有八种氨基酸。4 种带有脂肪烃侧链的氨基酸, 即丙氨酸、缬氨酸、亮氨酸和异亮氨酸; 两种含芳香环氨基酸: 苯丙氨酸和色氨酸; 一种含硫氨基酸即甲硫氨酸和一种亚氨基酸: 脯氨酸。其中甘氨酸介于极性和非极性氨基酸之间, 因其 α -碳原子上的氢受到其氨基和羧基的影响, 体现不出它的极性, 故常被列于非极性类。

2. 极性中性氨基酸

其特征是含有极性 R 侧链(如侧链上含有羟基或巯基、酰胺基等极性基团), 故有亲水性, 但在中性水溶液中不电离。这一组有 7 种氨基酸, 包括丝氨酸、苏氨酸、酪氨酸、天冬酰胺、谷氨酰胺、半胱氨酸。

3. 酸性氨基酸

其特征是它们的侧链含有羧基, 以解离出 H^+ 而具有酸性。这一组有两种氨基酸: 天冬氨酸和谷氨酸。

4. 碱性氨基酸

其特征是它们的侧链含有易接受 H^+ 的基团而具有碱性。这一组有三种氨基酸: 组氨酸、赖氨酸、精氨酸。

(三) 氨基酸的主要理化性质

1. 一般物理性质

氨基酸为无色晶体, 有很多熔点, 一般熔点超过 200°C , 个别超过 300°C , 原因是氨基酸在晶体时也以两性离子的形式存在, 类似于无机盐。许多氨基酸在达到熔点或接近熔点时, 或多或少地发生分解, 因此熔点不明显。氨基酸一般都溶于水, 而且水溶液具有较高的介电常数。氨基酸不溶或微溶于有机溶剂(乙醇、乙醚)。可以用乙醇沉淀氨基酸。氨基酸均可溶于酸、碱溶液中, 在配制氨基酸溶液时常用稀酸助溶。

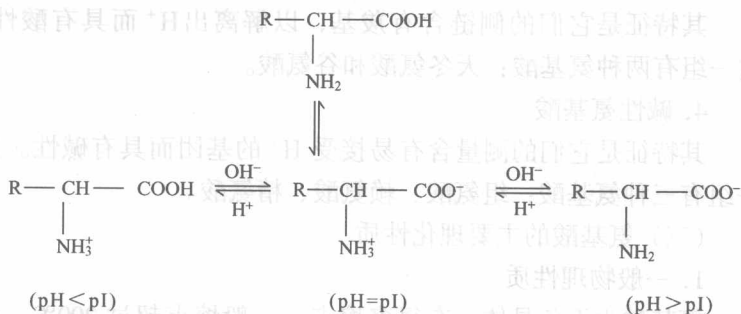
除甘氨酸外, α -氨基酸的 α -碳是一个手性碳原子, 因此, α -氨基酸都具有旋光性, 不同氨基酸的旋光方向并不一致。比

旋是 α -氨基酸的物理常数之一，它是鉴别各种氨基酸的一种根据。用测定比旋光度的方法还可以测定氨基酸的纯度。

氨基酸不吸收可见光，在远紫外光区有吸收。参与蛋白质组成的氨基酸中色氨酸、酪氨酸和苯丙氨酸在近紫外区有光吸收，蛋白质因含有这些氨基酸，所以也有紫外吸收能力，一般最大光吸收在 280nm 处，这是紫外吸收法定量蛋白质的依据。

2. 氨基酸的解离性质

所有氨基酸都含有碱性的氨基（或亚氨基），又含有酸性的羧基，故既有碱的性质，又有酸的性质，因此是两性电解质，具有两性游离性质；在不同 pH 值溶液中，可以带不同的电荷。在酸性溶液中羧基与质子（ H^+ ）结合而带正电荷，在碱性溶液中 NH_3^+ 上的 H^+ 与 OH^- 结合而带负电荷。氨基酸的解离过程如下：



由上式可见，当将氨基酸溶液的 pH 值调到某一特定的 pH 值时，其酸性基团所产生的负电荷等于碱性基团所产生的正电荷，即为氨基酸的兼性离子，分子呈电中性。使某氨基酸所带的正、负电荷数相等时溶液的 pH 值称为该氨基酸的等电点，通常用 pI 表示。各种氨基酸所含的氨基、羧基的数目不同，而且各种基团解离程度不同，因此不同的氨基酸有各自特定的等电点（见表 1-1）。

氨基酸的等电点及表现解离常数可以用它的酸碱滴定曲线的

实验方法求出。图 1-1 为丙氨酸的酸碱滴定曲线。

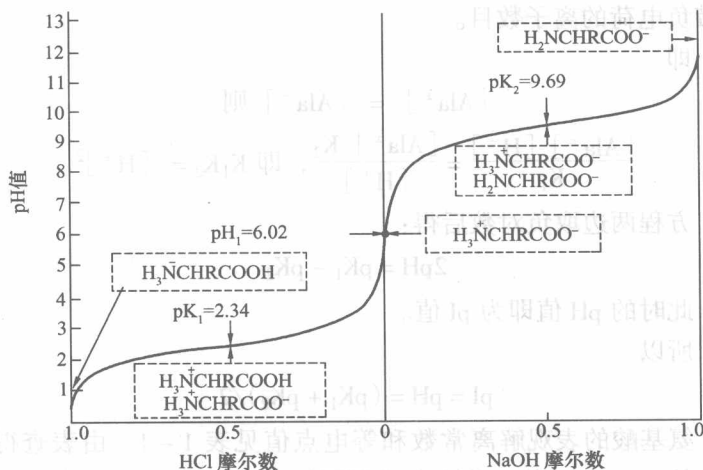
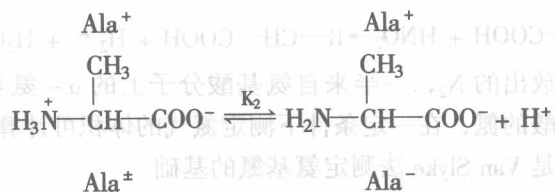
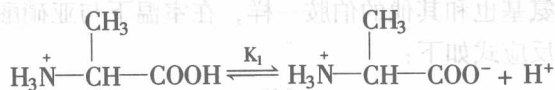


图 1-1 丙氨酸的酸碱滴定曲线

(虚线方框内表示在滴定过程中占优势的离子形式；R 代表丙氨酸的甲基)

以丙氨酸为例说明氨基酸的解离情况，它分步解离如下：



$$K_1 = \frac{[\text{Ala}^+][\text{H}^+]}{[\text{Ala}^+]} \quad [\text{Ala}^+] = \frac{[\text{Ala}^+][\text{H}^+]}{K_1}$$

$$K_2 = \frac{[\text{Ala}^-][\text{H}^+]}{[\text{Ala}^+]} \quad [\text{Ala}^-] = \frac{[\text{Ala}^+]}{[\text{H}^+]} K_2$$