

生物化学是 20 世纪的科学

生物化学是生命的化学,是利用化学的原理在分子水平上去解释生物学的科学。生物化学作为一门蓬勃向上的科学出现只是近一百年的事,可以说生物化学是 20 世纪的科学。在 19 世纪末和 20 世纪初,生物学家已经能够阐明一些生命的基本化学过程,但推动生物化学飞跃发展要归功于生物化学领域内的两个重要突破:酶的催化作用和核酸的遗传信息的阐明。

人们将酶的作用理解为生物反应催化剂的作用是在 19 世纪末和 20 世纪初的世纪之交,这是生物化学领域内的一个重要突破,普鲁士化学家 Emil Fischer 作出了很大的贡献。他研究了酵母酶对二糖水解的催化作用。Fischer 提出了在催化期间,酶和底物(反应物质)结合形成了酶-底物复合物,同时他认为只有结构合适的分子才能作为酶的底物。Fischer 把酶看作是一个刚性的模板,或是一把锁,而将底物看作是与之锁相配的钥匙。虽然这种理论不能圆满解释有些酶的作用机制,但这种酶作用的锁-钥理论仍是现代生物化学的中心信条。

Eduard Buchner 的研究工作使人们对酶有了进一步的认识。1897 年 Buchner 的实验证明,酵母细胞的提取液能够催化葡萄糖发酵生成酒精和二氧化碳,但是以前人们一直认为只有活细胞才能催化这样复杂的反应。实际上生命中的所有反应都是由酶催化的。

第二个突破发生在 Fischer 理论提出和 Buchner 实验 50 年之后,核酸作为信息分子的生物学作用被阐明。1944 年 Pswald Avery 等人从有毒的带有荚膜的肺炎双球菌提取出脱氧核糖核酸(DNA),然后与无毒的无荚膜的肺炎双球菌混合,结果无荚膜的肺炎双球菌被转化为有毒的有荚膜的菌种。这个实验首次提供了 DNA 是遗传物质的证据。1953 年,James D. Watson 和 Francis H. C. Crick 提出了 DNA 双螺旋模型。双螺旋模型使 Watson 和 Crick 立刻意识到 DNA 能够复制,就是说,遗传信息可以高保真地传给下一代。Watson 和 Crick 的工作被科学界公认为 20 世纪最伟大的科学发现之一。

许多化学元素是生命所需要的

人们经常遇到的大多数复杂的生物分子其实只是由少数几种元素组成的,大多数有机体的 97% 以上的质量是由氧、碳、氢、氮、磷和硫 6 种元素组成的,当然在不同的有机物中,这 6 种元素变化很大。碳是细胞的固体材料中最常见的元素。碳原子之间可以成键,碳原子也可以与氢、氮、氧和硫等其他元素的原子成键。由于碳原子形成的共价键数目可以达到 4 个,所以可以生成多种多样的化合物。在生物化学中含碳的化合物都被看作是有机化合物(但像 CO_2 等少数化合物除外)。表 0.1 给出了生物化学中常见的一些典型有机化合物、反应官能团和常见的化学键类型。表中介绍的这些化合物、官能团和化学键将反复出现在本书中,读者应熟记它们的结构和反应特性。

表 0.1 生物化学中常见的一些有机化合物、反应官能团和常见的键的类型

(a) 有机化合物				
$\text{R}-\text{OH}$ 醇	$\text{R}-\overset{\text{O}}{\parallel}{\text{C}}-\text{H}$ 醛	$\text{R}-\overset{\text{O}}{\parallel}{\text{C}}-\text{R}_1$ 酮	$\text{R}-\overset{\text{O}}{\parallel}{\text{C}}-\text{OH}$ 羧酸 ¹	
$\text{R}-\text{SH}$ 巯基化合物	$\text{R}-\text{NH}_2$ 一级	$\text{R}-\overset{\text{R}_1}{\underset{ }{\text{N}}}\text{H}$ 二级	$\text{R}-\overset{\text{R}_1}{\underset{ }{\text{N}}}-\text{R}_2$ 三级	胺 ²
(b) 官能团				
$-\text{OH}$ 羟基	$-\overset{\text{O}}{\parallel}{\text{C}}-\text{R}$ 酰基	$-\overset{\text{O}}{\parallel}{\text{C}}-$ 酮基	$-\overset{\text{O}}{\parallel}{\text{C}}-\text{O}^-$ 羧基	
$-\text{SH}$ 巯基	$-\text{NH}_2$ 或 $-\overset{+}{\text{N}}\text{H}_3$ 氨基	$-\text{O}-\overset{\text{O}}{\parallel}{\text{P}}(\text{O}^-)-\text{O}^-$ 磷酸基	$-\overset{\text{O}}{\parallel}{\text{P}}(\text{O}^-)-\text{O}-$ 磷酰基	
(c) 化学键				
$\begin{array}{c} \\ \text{---C---O---C---} \\ \end{array}$ 酯键	$\begin{array}{c} \quad \\ \text{---C---O---C---} \\ \quad \end{array}$ 醚键	$\begin{array}{c} \text{O} \\ \\ \text{---N---C---} \\ \end{array}$ 酰胺键		
$\begin{array}{c} \\ \text{---C---O---P---O---} \\ \quad \\ \quad \text{O}^- \end{array}$ 磷酸酯键	$\begin{array}{c} \text{O} \quad \text{O} \\ \quad \\ \text{---O---P---O---P---O---} \\ \quad \quad \quad \\ \text{O}^- \quad \text{O}^- \quad \text{O}^- \quad \text{O}^- \end{array}$ 磷酸酐键			

1. 在大多数条件下，羧酸表示为解离状态： $\text{R}-\overset{\text{O}}{\parallel}{\text{C}}-\text{O}^-$ ；
2. 胺可以表示为质子化形式： $\text{R}-\overset{+}{\text{N}}\text{H}_3$ ， $\text{R}-\overset{+}{\text{N}}\text{H}_2$ 和 $\text{R}-\overset{+}{\text{N}}\text{H}-\text{R}_2$ 。

在活的有机体中，除了上述 6 种元素以外，还发现了另外二十几种元素，例如钙、镁、钠、氯、钾和铁等元素。有些元素在机体中起着重要的作用，例如有些酶的催化作用就需要金属离子作为辅助因子。

许多重要的生物分子是聚合物

从某种意义上说，生物化学是生物聚合物的化学，这些聚合物是由许多小分子彼此连

接形成的。具有化学知识的读者可能对这些小分子比较熟悉，但要深入到生物化学中，就要非常熟悉分子量从几千到上百万的常见的生物大分子。除了水之外，细胞内最丰富的化合物是生物聚合物，丰富程度的高低依次为蛋白质、核酸、多糖和脂。

形成聚合物的每个构件分子称为单体，由单体形成聚合物需要连续多步的缩合反应，整合到聚合物的单体常称为残基。有的聚合物是由一种单体聚合形成的，像一些多糖是由单一葡萄糖残基多次重复的产物，而蛋白质和核酸则是由不同残基按照特定顺序聚合的产物，残基的序列决定着它的功能。

生物聚合物所具有的特性明显地不同于它的单体特性。这一特点导致生命的组织层次的形成。无论对单细胞，还是多细胞生物，细胞都是生命的基本单位。图 0.1 给出了细胞的分子组织层次。可以看出，活细胞是由无生命的分子组成的，是这些无生命的分子组装、相互作用赋予细胞以生命。

蛋白质和核酸是两种类型的大的聚合物，多次出现在本书的章节中，说明了它们在组织层次中的重要性。蛋白质是能够折叠成三维形状（构象）的线性聚合物，它的活性取决于它的构象。一种酶能够选择某一特定的底物，催化底物生成产物的能力是不能通过它的构成成分的化学特性来解释的，只能通过酶的有组织的三维形状来解释。与蛋白质类似，DNA 是贮存遗传信息的聚合物。DNA 的四个核苷酸成分单独没有承载信息的能力。然而，当核苷酸连接成线性的 DNA 聚合物时，DNA 的序列就能够编码和指导细胞内所有蛋白质的合成。

蛋白质是氨基酸的聚合物

我们将了解可以整合到所有细胞蛋白质中的 20 种标准氨基酸。氨基酸可以由简单的前体分子合成，或以营养物质被摄入有机体。所有的氨基酸都至少含有两个功能基团——出现在 α 碳上的一个氨基和一个羧基。氨基酸之间的差别主要表现在侧链上。最简单的氨基酸是甘氨酸，它的侧链是一个氢原子。一个氨基酸的羧基可以与另一个氨基酸的氨基缩合形成肽，氨基酸之间是通过称之为肽键的酰胺键共价连接的。而蛋白质就是氨基酸通过肽键共价连接的聚合物。自然界中发现的各种各样的蛋白质，它们的氨基酸组成和氨基酸的排列顺序都是不同的。蛋白质的功能取决于它的三维结构，而蛋白质的三维结构主要是由它的氨基酸序列确定的，当然，氨基酸的序列最终又是由基因编码的。

多糖是单糖的聚合物

糖主要是由碳、氢和氧组成的。糖包括单糖、多糖和其他的糖衍生物。所有的单糖都含有几个羟基，因此也称为聚醇。最常见的单糖是五碳糖或六碳糖。葡萄糖是主要的一

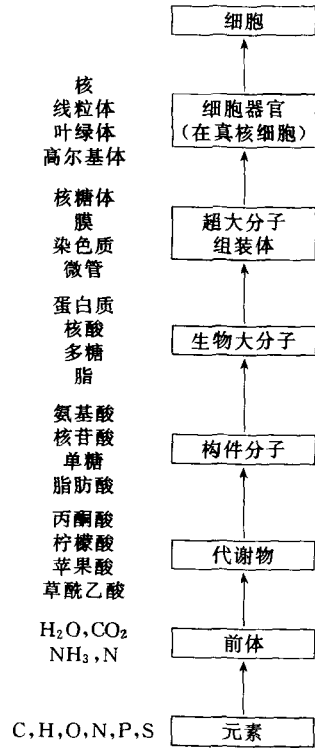


图 0.1 细胞的分子组织层次

种六碳糖，它是细胞的营养物质，是贮存多糖糖原和淀粉的单体。纤维素是植物细胞壁的结构多糖，也是葡萄糖的聚合物。多糖是单糖通过形成糖苷键缩合形成的。核糖是最常见的五碳糖，它是核糖核酸（RNA）的糖成分，而出现在脱氧核糖核酸（DNA）中的是 2-脱氧核糖。就像我们将看到的那样，核糖和脱氧核糖是核苷酸的组成成分，而核苷酸是核酸的构件分子。

核酸是核苷酸的聚合物

在核酸一章(第 7 章)我们将会看到核苷酸含有一个核糖或脱氧核糖、一个含氮的杂环碱基和至少一个磷酸基团。核苷酸中的碱基是嘌呤和嘧啶碱基。主要碱基有腺嘌呤(A)、鸟嘌呤(G)、胞嘧啶(C)、胸腺嘧啶(T)和尿嘧啶(U)。

在核酸分子中，核苷酸之间通过磷酸二酯键共价连接。共价连接的核苷酸一般称为核苷酸残基。核酸含有许多核苷酸残基。在 DNA 分子中，两股核酸链的碱基以配对的方式相互作用，A 与 T 配对，而 G 与 C 配对，整个 DNA 分子形成一个螺旋结构。特定的碱基序列负载着遗传信息。

有些核酸分子很大，例如人的一些染色体含有一个亿以上核苷酸残基的 DNA 分子。但细菌含有的基因比哺乳动物少，通常含有小的、环状的 DNA 分子。

RNA 是由一个 DNA 模板合成的单链聚核苷酸。主要存在着三种类型的 RNA，包括直接来自 DNA 的信息转移给蛋白质的信使核糖核酸（mRNA）在 mRNA 中，每三个核苷酸残基组成一个密码，编码着特定的氨基酸。蛋白质是通过解读 mRNA 上的密码合成的。转移核糖核酸（tRNA）是比较小的用于蛋白质合成的分子。而核糖体核糖核酸（rRNA）是核糖体的主要成分，核糖体是 RNA 和蛋白质的复合体，它是蛋白质合成的场所。

脂是形式多种多样的分子

脂是一类定义为不溶于水的有机分子化合物。某些脂是贮能分子，某些脂是膜的结构成分，还有些脂参与细胞内和细胞之间的通讯。最简单的脂是脂肪酸，它是带有羧基的长的碳氢链。脂肪酸常出现在三脂酰甘油（或脂肪）和甘油磷脂的大分子中。三脂酰甘油是哺乳动物的最丰富的一类脂，它含有与一分子甘油成酯的三个脂肪酸成分。甘油磷脂是由与两个脂肪酸成酯的甘油-3-磷酸组成，一般都含有一个与磷酸连接的极性头（例如乙醇胺）。磷脂是生物膜的主要成分。

膜脂一般都含有一个极性的、亲水的头部和一个非极性的疏水尾巴。在水溶性的介质中，膜脂的疏水尾巴相互靠近形成两个片层，称为脂双层。脂双层构成了所有生物膜结构的基础。生物膜将细胞或细胞内的区室与周围的环境隔离开。大多数生物膜都含有镶嵌在膜内或附着在膜上的蛋白质。某些膜蛋白可以用作营养物质进入或废物排出的通道，而有些膜蛋白可以催化发生在膜表面的特异反应。

热力学和生物力能学的知识在生物化学中很重要

生物化学是在包括生物分子合成和降解的所有的反应在内的化学水平上研究生命的科学。有机体需要能量来维持它们的活力、生长和繁殖。对地球上的所有生物来说，太阳是能量的最终来源。植物通过光合作用捕获太阳能用于合成糖，而许多动物又从植物中获取营养物。所有生物都是以自由能的形式从营养物中获得能量用于驱动生命过程，多

余的能量又以热的形式释放到环境中。积累的多余的能量增加了环境的无序性（熵），所以活的有机体都是在破坏环境的有序性来维持它自身的组织性（有序性）的。

研究生物化学反应期间分子的能量和能量转换的科学称为生物力能学。生物力能学是热力学领域的一部分，是涉及能量转换的物理学的的一个分支。应用于非生命系统的基本的热力学原理也可以应用于生命的化学。在生物化学中，最有用的热力学变量是反应的自由能变化（ ΔG ）。从 ΔG 可以看出一个反应是否能自发进行，也能计算出一个反应的平衡常数。 ΔG 含有两个变量 焓（ H ）和熵（ S ）。这些热力学术语在全书中经常遇到，并将在第 8 章代谢导论中详细说明。对这些热力学术语的了解有助于理解和掌握生物化学反应的过程和机制。

生物化学的学习是一个渐进的过程

这本生物化学教材是一本逻辑性很强的教材，各个章节的编排次序都是一环扣一环的，就是说，前一章的内容是后一章的基础。例如，一般先介绍生物化学单体的结构，然后紧接着讨论相应的聚合物的结构、功能和生物合成。所以生物化学的学习过程需要跨越几个台阶。掌握每一类生物分子的结构可以说是生物化学学习所要达到的第一个台阶。化学结构是生物化学的词汇，大量词汇的积累只有通过反复记忆。分子结构的记忆在整个生物化学课程学习过程中会使你受益匪浅。例如开始学习的 20 种氨基酸的结构在以后的许多章节中都会遇到，掌握了它们的结构，会使你易于理解各个氨基酸的理化特性、形成聚合物的特点、反应的机制以及生物合成和降解的一些规律。

代谢的学习是生物化学中的第二个台阶。与分子的静态结构相比，细胞内的化学转换处于动态之中。几乎所有代谢中的反应，无论是生物合成，还是降解，都是由酶催化的。物质代谢总是伴随着能量代谢，在燃料分子进行分解代谢降解为中间代谢物的同时，释放出大量的能量，有些中间代谢物又是合成的前体分子。在代谢途径中首先要了解和熟悉的是碳代谢的中心代谢途径——糖酵解和柠檬酸循环，这一途径是糖、蛋白质和脂等生物分子代谢的必经之路。

生物化学学习的最后一个台阶是了解遗传信息分子的复制、转录、翻译及其调控。这一部分学习起来有一定的难度，尤其是基因调控内容。主要是要了解 DNA 复制的一些基本规律、转录后的 RNA 加工过程和遗传信息翻译成蛋白质的原则。

在即将到来的 21 世纪，许多充满神秘色彩的生命现象会逐渐为人们所破译，许多折磨人类的遗传疾病会被人们攻克。随着生命科学的蓬勃发展，生命的奥秘会逐渐为人们所了解和认识。希望读者利用好本教材，打下雄厚的生物化学基础，为生命科学及其相关科学的发展作出贡献。

蛋白质(protein)是生物体内的主要的生物分子,因生物种类的不同,其蛋白质的种类和含量有很大的差别,例如人体内大约含有 30 万种蛋白质,一个大肠杆菌的蛋白质含量虽少,但也含有 1 000 种以上。整个生物界约有 $10^{10} \sim 10^{12}$ 种蛋白质。但无论是人体内的蛋白质,还是大肠杆菌中的蛋白质,都主要是由 20 种氨基酸(amino acid)构成的,即这 20 种氨基酸是生物界通用的或是标准的氨基酸。

1.1 蛋白质是由 20 种不同的氨基酸构成的

蛋白质是由氨基酸构成的聚合物,所有生物都是利用这 20 种氨基酸作为构件组装成各种蛋白质分子的。尽管氨基酸的种类有限,但由于氨基酸在蛋白质中连接的次序以及氨基酸数目的不同,可以组装成几乎无限的不同种类的蛋白质。

20 种氨基酸都被称为 α -氨基酸,因为 α -氨基酸分子中的 α -碳(分子中第二个碳, C_α)结合着一个氨基和一个酸性的羧基,此外, C_α 还结合着一个 H 原子和一个侧链基团(用 R 表示)。每一种氨基酸的 R 都是不同的,侧链上的碳依次按字母命名为 β, γ, δ 和 ϵ 碳,分别指的是第三、四、五和六位碳:

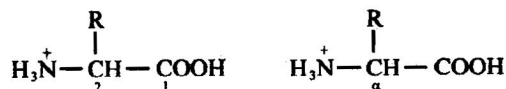
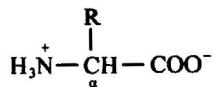


表 1.1 给出了 20 种标准氨基酸的英文名称以及三字母缩写和单字母表示符号。在表示蛋白质的组成和氨基酸序列时常常使用这些氨基酸的简写符号。

在中性 pH(生理条件)下,氨基质子化($-\text{NH}_3^+$),而羧基离子化($-\text{COO}^-$)。 α -羧基的 pK_a 值的范围是 1.8~2.5,而 α -氨基的 pK_a 值的范围是 8.7~10.7,因此在生理 pH(7.4)下,氨基酸是个兼性离子,或称之偶极离子:



下面我们将看到有些侧链也是可以离子化的。如果与 C_α 连接的侧链 R 不是 H 原子,则 C_α 就结合了 4 种不同的基团,在这种情况下, C_α 是个不对称碳,含有一个不对称碳的氨基酸就存在着两种不能叠合的镜像立体异构体。20 种氨基酸中有 19 种标准氨基酸存在着立体异构体。甘氨酸是个例外,因为它的 R 基团是一个 H,由于 C_α 上连有两个 H,所以甘氨酸就不存在立体异构体。立体异构体是指分子式相同,但分子中的原子在空间排列(或称为构型)不同的化合物。如果要改变构型,需要破坏一个或更多的化学键。

表 1.1 20 种标准氨基酸的中英文名称及简写符号

中文名称	英文名称	三字母缩写	单字母符号
甘氨酸	Glycine	Gly	G
丙氨酸	Alanine	Ala	A
缬氨酸	Valine	Val	V
亮氨酸	Leucine	Leu	L
异亮氨酸	Isoleucine	Ile	I
脯氨酸	Proline	Pro	P
苯丙氨酸	Phenylalanine	Phe	F
酪氨酸	Tyrosine	Tyr	Y
色氨酸	Tryptophan	Trp	W
丝氨酸	Serine	Ser	S
苏氨酸	Threonine	Thr	T
半胱氨酸	Cystine	Cys	C
蛋氨酸	Methionine	Met	M
天冬酰胺	Asparagine	Asn	N
谷氨酰胺	Glutamine	Gln	Q
天冬氨酸	Aspartic acid	Asp	D
谷氨酸	Glutamic acid	Glu	E
赖氨酸	Lysine	Lys	K
精氨酸	Arginine	Arg	R
组氨酸	Histidine	His	H

按照惯例,将 $\alpha\text{-COO}^-$ 画在顶端,垂直地画一个氨基酸,然后与立体化学参考化合物甘油醛-3-磷酸比较, α -氨基位于 α -碳左边的是 L-异构体,位于右边的为 D-异构体,氨基酸的一对镜像异构体分别为 D 型和 L 型异构体。到目前为止,所发现的游离的和蛋白质温和水解得到的氨基酸绝大多数是 L 型氨基酸;D 型氨基酸主要存在于微生物中,图 1.1 给出了丝氨酸的两种立体异构体。

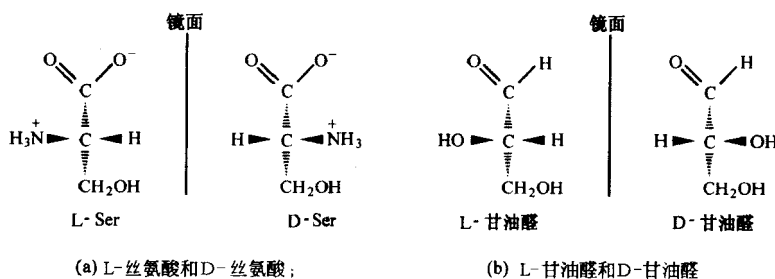


图 1.1 丝氨酸立体异构体与 L-和 D-甘油醛构型的立体关系

1.1.1 20种氨基酸可以按其侧链分类

图 1.2 以 Fischer 投影式给出了通常在蛋白质中发现的 20 种氨基酸的结构式。Fischer 投影式中,处于手性中心的水平键表示朝向读者,而垂直键表示背离读者。从图 1.2 中可

以看到,20种氨基酸的侧链有很大区别,有些侧链是非极性的,表现出疏水性;而其他一些侧链是极性的或是在中性 pH 下离子化的,表现出亲水性。由于单个氨基酸是兼性离子,所以是水溶性的。但值得注意的是,当氨基酸通过聚合形成多肽或蛋白质后,氨基酸中的 α -氨基和 α -羧基上的离子电荷就失去了,结果侧链的非极性、极性或离子特性对于一个蛋白质的整个三维构象有很大的影响。20种氨基酸按照它们的侧链基团可以分为6种脂肪族、3种芳香族、2种含硫、2种醇类、3种碱性、2种酸性和2种酰胺氨基酸。

(1) R为脂肪族基团的氨基酸

甘氨酸的结构在20种氨基酸中最简单,因为它的侧链是个氢原子,所以甘氨酸的 α -碳不是手性碳。 α -碳上的2个H赋予分子一点疏水特性。甘氨酸在许多蛋白质构象中起着独一无二的作用,因为它的侧链很小,可以填进不能容纳别的氨基酸的缝隙里。

丙氨酸、缬氨酸、亮氨酸和异亮氨酸4种氨基酸都带有饱和的脂肪烃链。丙氨酸(α -氨基丙酸)的侧链是一个简单的甲基;缬氨酸(α -氨基异戊酸)的侧链是一个含有支链的3碳侧链;亮氨酸(α -氨基异己酸)和异亮氨酸(α -氨基- β -甲基戊酸)都含有带支链的4碳侧链。要注意的是,异亮氨酸分子中的 α -碳和 β -碳都是不对称碳,所以异亮氨酸存在着4种可能的立体异构体:L-异亮氨酸,D-异亮氨酸,L-别构异亮氨酸和D-别构异亮氨酸(别构也是一种异构形式)。通常出现在蛋白质中的L-异亮氨酸只是4个异构体中的一种。尽管丙氨酸、缬氨酸、亮氨酸和异亮氨酸没有反应功能基团,但这些氨基酸在建立和维持蛋白质的三维结构中起着重要的作用,它们倾向于聚集以避开水,因为缬氨酸、亮氨酸和异亮氨酸的侧链是高度疏水的。

脯氨酸(β -吡咯烷基- α -羧酸)明显地不同于其他19种氨基酸,它有一个环形的饱和烃侧链,结合在 α -氨基的氮和 α -碳上。所以严格地讲,脯氨酸是一个亚氨基酸,因为它含有一个二级氨基,而不是一级氨基。不过,在蛋白质中发现的脯氨酸也像其他氨基酸一样,它的 α -碳也是手性碳。脯氨酸的杂环吡咯烷环限制多肽的几何构型,有时会在多肽链中引进一个转折的变化。

(2) R为芳香族基团的氨基酸

苯丙氨酸、酪氨酸和色氨酸都含有芳香族基团的侧链。苯丙氨酸(α -氨基- β -苯丙酸)是一个含有苯基的氨基酸。酪氨酸(α -氨基- β -对羟基苯丙酸)的结构类似于苯丙氨酸,称为对羟基苯丙氨酸,即酪氨酸是个带有酚基的氨基酸。酚是个比较弱的酸,所以酪氨酸侧链的 pK_a 值为10.5。色氨酸(β -吲哚- α -氨基丙酸)的侧链带有一个双环的吲哚基,所以色氨酸又称为吲哚丙氨酸。由于酪氨酸和色氨酸的侧链带有极性基团,所以它们的疏水性不像苯丙氨酸那样强。这三种芳香族氨基酸都吸收紫外光,在中性 pH 条件下,色氨酸和酪氨酸的紫外吸收峰在280nm,而苯丙氨酸的在260nm处。大多数蛋白质中都含有色氨酸和酪氨酸或其中的一种,所以溶液中280nm吸收的测量常用来估算蛋白质的浓度。

(3) R为含硫基团的氨基酸

蛋氨酸和半胱氨酸是两个含硫氨基酸。蛋氨酸(α -氨基- γ -甲硫基丁酸)是疏水氨基酸,它的侧链上带有一个非极性的甲基硫醚基,不过,其中的硫原子是亲核的。半胱氨酸(α -氨基- β -巯基丙酸)侧链上含有一个巯基($-SH$),所以又称为巯基丙氨酸。虽然半胱氨酸

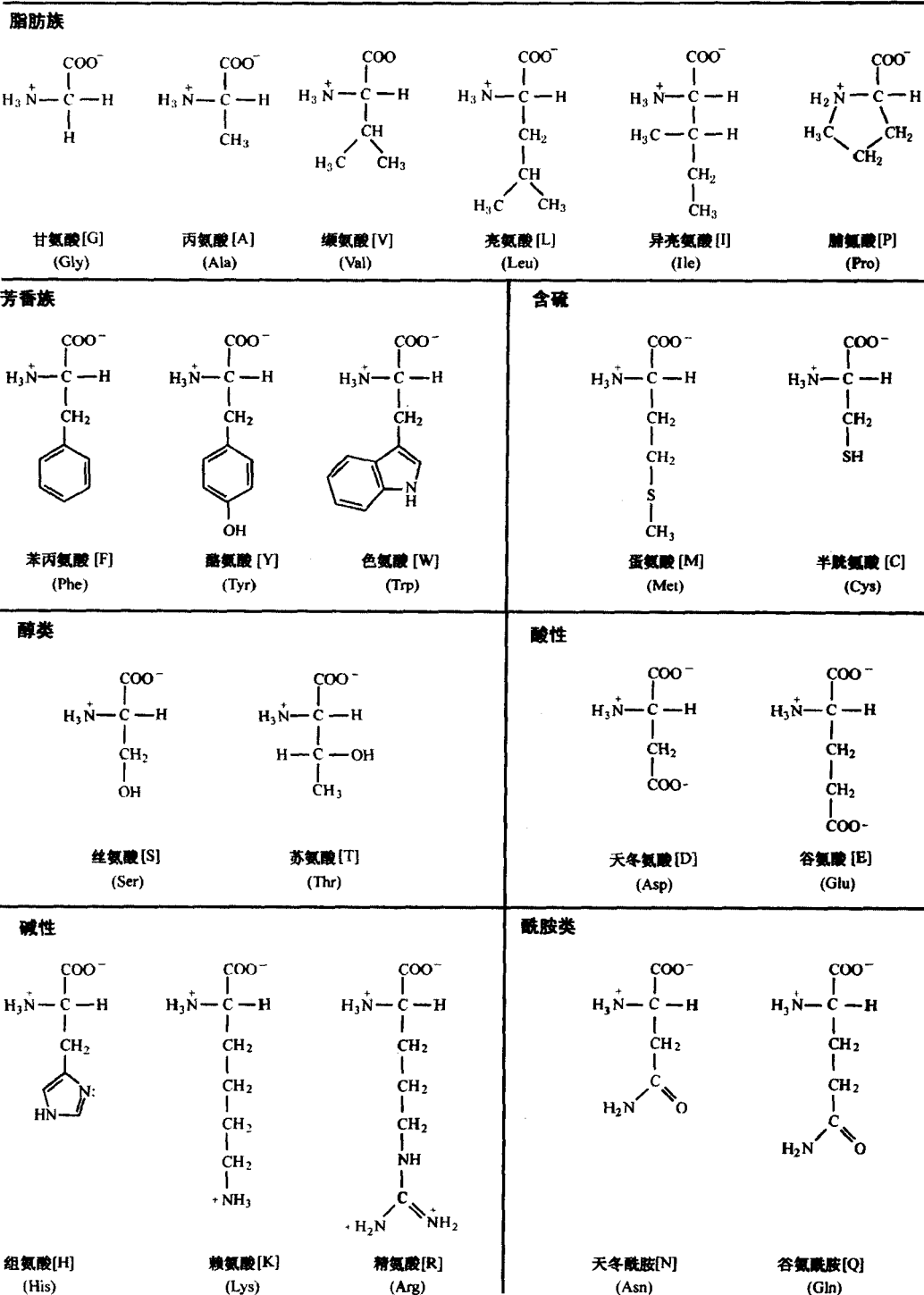


图 1.2 用 Fischer 投影式给出的 20 种氨基酸的结构

酸的侧链显得有点疏水性，但—SH 也是一个高反应性的基团。因为硫原子是个可极化的原子，巯基能与氧和氮形成弱的氢键。半胱氨酸的巯基像弱酸 ($pK_a = 8.4$)，它可以失去质子。当某些蛋白质水解时，会分离出一种称为胱氨酸的化合物。胱氨酸是由两个半胱氨酸通过一个二硫键连接形成的，二硫键是两个半胱氨酸的巯基氧化形成的。二硫键，或称为二硫桥，在稳定某些蛋白质的三维结构中起着重要的作用，蛋白质中二硫桥通常都是处于多肽链的不同位置的半胱氨酸通过共价交联形成的。

(4) R 为含醇基基团的氨基酸

丝氨酸和苏氨酸是两个侧链含有 β -羟基的不带电荷的氨基酸。苏氨酸 (α -氨基- β -羟基丙酸) 和丝氨酸 (α -氨基- β -羟基丁酸) 的羟基与酪氨酸的酚基不同，它具有一级和二级醇的弱的离子化特性。尽管丝氨酸的羟甲基 ($-\text{CH}_2\text{OH}$) 在水溶液中不显示离子化，但这个醇可以参与很多酶的活性部位反应，就像已被离子化了一样。注意，苏氨酸也像异亮氨酸一样，具有两个手性碳原子，即 α -碳和 β -碳原子，通常出现在蛋白质中的 L-苏氨酸只是 4 种立体异构体中的一种。

(5) R 为碱性基团的氨基酸

有 5 种氨基酸在 pH7 时它们的侧链基团带有电荷，其中有 3 种是带正电荷的碱性 R 基团，另外 2 种是带负电荷的酸性 R 基团。组氨酸、赖氨酸和精氨酸都带有亲水性的含氮碱基基团。组氨酸 (α -氨基- β -咪唑基丙酸) 的侧链有一个咪唑环，所以又称为咪唑丙氨酸。咪唑基是可以离子化的 ($pK_a = 6.0$)。赖氨酸 (α, ϵ -二氨基己酸) 是一个双氨基酸，含有 α -和 ϵ -氨基。在中性 pH, ϵ -氨基是以碱性氮离子 ($-\text{CH}_2-\text{NH}_3^+$) 形式存在的，在蛋白质中通常带正电荷。精氨酸 (ζ -氨基- δ -胍基戊酸) 是 20 种氨基酸中碱性最强的氨基酸，它的侧链胍基离子的 pK_o 值最高 ($pK_a = 12.5$)。

(6) R 为酸性基团的氨基酸

天冬氨酸 (α -氨基丁二酸) 和谷氨酸 (α -氨基戊二酸) 是二羧基氨基酸，除了含有 α -羧基外，天冬氨酸还含有 β -羧基，谷氨酸还含有 γ -羧基。由于两个氨基酸的侧链在 pH7 时都离子化了，所以在蛋白质中是带负电荷的。这两个氨基酸经常出现在蛋白质分子表面上。谷氨酸的单钠盐是调味品味精。

(7) R 为含酰胺基团的氨基酸

天冬酰胺 (β -氨基- β -羧基丙酰胺) 和谷氨酰胺 (γ -氨基- γ -羧基丁酰胺) 分别是天冬氨酸和谷氨酸的酰胺化产物。尽管这两个氨基酸的侧链是不带电荷的，但它们的极性很强，也经常出现在蛋白质分子的表面，它们可以和水相互作用。另外，这两种氨基酸的酰胺基可以与其他极性氨基酸的侧链上的原子形成氢键。

以上氨基酸完全是根据侧链的性质分类的，从营养学角度又可将氨基酸分为必需和非必需氨基酸。必需氨基酸包括赖氨酸、苯丙氨酸、蛋氨酸、亮氨酸、异亮氨酸、缬氨酸、苏氨酸、色氨酸、组氨酸和精氨酸，其中组氨酸和精氨酸，人虽然能合成，但效率低，尤其在婴幼儿时期，需要由外界供给。非必需氨基酸是其余的 10 种氨基酸：甘氨酸、丝氨酸、半胱氨酸、酪氨酸、谷氨酸、谷氨酰胺、天冬氨酸、天冬酰胺、脯氨酸和丙氨酸。

必需和非必需氨基酸都是人体所需要的氨基酸，只不过必需氨基酸人自己不能合成或合成的量很少，不能满足人体的需要，必须由食物供给；非必需氨基酸是人体自己可以

合成,不必由食物供给的氨基酸。

1.1.2 氨基酸的离子状态取决于环境的 pH

由于氨基酸含有酸性的 α -羧基和碱性的 α -氨基,有些氨基酸的侧链还含有可解离基团,所以每种氨基酸都有 2 个或 3 个 pK_a 值,在给定的 pH 下,氨基酸常带有不同的净电荷。

表 1.2 给出了 25℃ 下游离氨基酸的酸性和碱性基团的 pK_a 值和氨基酸的等电点 (pI)。表中组氨酸的侧链基团咪唑基的 pK_a 值是 6.0,在生理 pH (pH7 左右) 有明显的缓冲作用,所以组氨酸是惟一的一个在生理条件下具有缓冲作用的氨基酸。

表 1.2 氨基酸的 pK_a 和 pI 值

氨基酸	分子量	pK_a			pI	蛋白质中出现的几率*
		α -COOH	α -NH ₃ ⁺	R 基团		
甘氨酸	75	2.34	9.60		5.97	7.5%
丙氨酸	89	2.34	9.69		6.01	9.0%
缬氨酸	117	2.32	9.62		5.97	6.9%
亮氨酸	131	2.36	9.60		5.98	7.5%
异亮氨酸	131	2.36	9.68		6.02	4.6%
脯氨酸	115	1.99	10.96		6.48	4.6%
苯丙氨酸	165	1.83	9.13		5.48	3.5%
酪氨酸	181	2.20	9.11	10.07	5.66	3.5%
色氨酸	204	2.38	9.39		5.89	1.1%
丝氨酸	105	2.21	9.15		5.68	7.1%
苏氨酸	119	2.11	9.62		5.87	6.0%
半胱氨酸	121	1.96	8.18	10.28	5.05	2.8%
蛋氨酸	149	2.28	9.21		5.74	1.7%
天冬酰胺	132	2.02	8.80		5.41	4.4%
谷氨酰胺	146	2.17	9.13		5.65	3.9%
天冬氨酸	133	1.88	9.60	3.65	2.77	5.5%
谷氨酸	147	2.19	9.67	4.25	3.22	6.2%
赖氨酸	146	2.18	8.95	10.53	9.74	7.0%
精氨酸	174	2.17	9.04	12.48	10.76	4.7%
组氨酸	155	1.82	9.17	6.00	7.59	2.1%

* 在 200 多种蛋白质中出现的平均几率。

氨基酸的 α -COOH 基是个弱酸,我们可以利用 Henderson-Hasselbalch 方程计算任何 pH 下离子化基团的比例:

$$pH = pK_a + \lg \frac{[\text{共轭碱}]}{[\text{弱酸}]}$$

从表 1.2 可以看出,游离氨基酸的 α -COOH 的 pK_a 值的范围是 1.8 至 2.5。这些值都比典型的羧酸低 (例如乙酸的 $pK_a = 4.8$),这主要是邻近的一 NH_3^+ 基团的诱导影响的结果。对于一个典型的氨基酸,它的 α -COOH 的 pK_a 值大约是 2,可以通过下式计算出它的共

轭碱（羧酸盐离子，或盐）对弱酸（羧酸，或质子化的形式）在中性 pH 的比例：

$$7.0 = 2.0 + \lg \frac{[\text{RCOO}^-]}{[\text{RCOOH}]}$$

羧酸盐对羧酸的比是 100 000 : 1，所以在中性 pH 条件下，占优势的成分是羧酸盐离子。

我们看看游离氨基酸的 α -氨基的离子化状况。 α -氨基的 pK_a 值的范围是 8.7 ~ 10.4。对于一个典型的氨基酸，它的 α -氨基的 pK_a 值大约为 10.0，它的共轭碱（游离的胺 $-\text{NH}_2$ ）对共轭酸（质子化的胺 $-\text{NH}_3^+$ ）的比，在 pH7 时为 1 : 1 000。计算表明，在中性 pH 条件下，游离氨基酸主要是以兼性离子形式存在，所以将一个氨基酸描述成带有 $-\text{COOH}$ 和 $-\text{NH}_2$ 基团的形式是不合适的，因为不存在这样一种 pH，即在这一 pH 下， $-\text{COOH}$ 和 $-\text{NH}_2$ 都占优势。

脯氨酸的氨基 ($pK_a = 10.6$) 在中性 pH 条件下也是质子化的碱基形式，虽然侧链与 α -氨基结合，脯氨酸在 pH7 时仍是一种兼性离子。

对于侧链含有可解离基团的氨基酸，例如酸性和碱性氨基酸，在充分质子化后可以看作是一个三元弱酸。图 1.3 给出了组氨酸侧链的咪唑离子的去质子过程。在 pH7 条件下，咪唑（共轭碱）对咪唑离子（共轭酸）的比是 10 : 1。在接近生理 pH 条件下，侧链质子化和呈中性的两种形式组氨酸的浓度都很高。蛋白质中组氨酸侧链可以是质子化的，也可以是去质子型的，这主要取决于它在蛋白质内所处的环境。组氨酸侧链的这一特性使得它可以在酶催化部位进行质子转移。

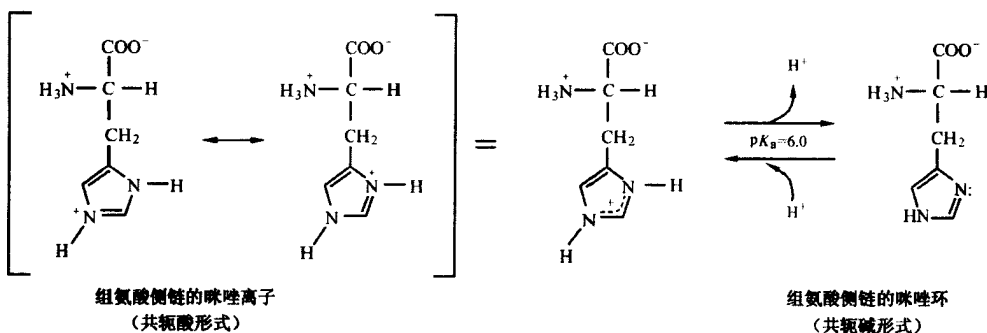


图 1.3 组氨酸侧链的咪唑环上质子的解离

通过氨基酸的滴定曲线可以确定氨基酸的各个解离基团的 pK_a 值。图 1.4 给出了丙氨酸和组氨酸的滴定曲线。丙氨酸有两个可解离基团， $\alpha\text{-COOH}$ 和 $\alpha\text{-NH}_3^+$ ，它们的 pK_a 值分别是 2.4 和 9.9，每一个值都是位于滴定曲线缓冲区的中心。组氨酸有 3 个可解离基团，即 $\alpha\text{-COOH}$ ， $\alpha\text{-NH}_3^+$ 和侧链基团咪唑基，它们的 pK_a 值分别是 1.8, 6.0 和 9.3，也都是处于缓冲区的中心。

pK_a 值等于一个可解离基团处于它的滴定中点时的 pH。换言之，就是在这个 pH 条件下，弱酸的浓度刚好等于它的共轭碱的浓度。人们可以推测出，丙氨酸分子上的净电荷在 pH2.4 时平均为 +0.5，而在 pH9.9 时是 -0.5。在 pH2.4 和 pH9.9 的中间，即在 pH6.15 时，溶液中的丙氨酸分子的净电荷为零。因此 pH6.15 称为丙氨酸的等电点

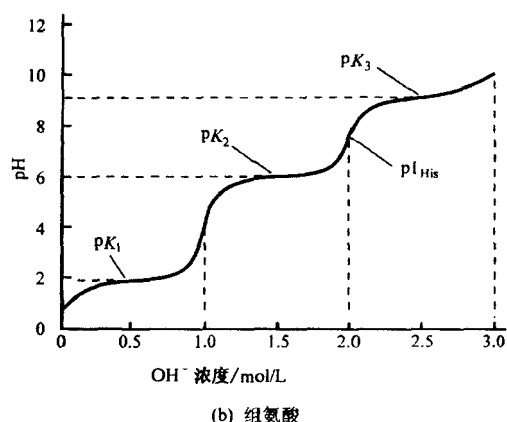
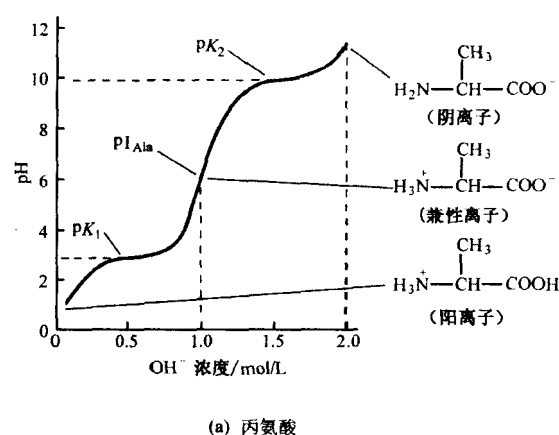


图 1.4 丙氨酸和组氨酸的滴定曲线

(isoelectric point, pI), 或等电 pH 。假如将丙氨酸置于一个电场中, 当丙氨酸所处溶液的 pH 比它的 pI 低时, 丙氨酸就带有净的正电荷 (换言之, 阳离子型的丙氨酸占优势), 将向电场的负极迁移; 而当 pH 高于 pI 时, 丙氨酸带负电荷, 将向电场的阳极迁移; 如果处于 pI ($pH=6.15$), 丙氨酸不会向任一方向迁移。氨基酸等电点的计算是很简单的, 像丙氨酸那样, 只有两个可解离基团的氨基酸的 pI 值, 就是该氨基酸的两个 pK_a 值的算术平均值。然而, 要求出像组氨酸那样有 3 个可解离基团的氨基酸的等电点, 首先要估算该氨基酸处于 3 个 pK_a 中的任一个时的平均净电荷。以组氨酸为例, 当它处于 $pH 1.8$ 时, 净电荷为 $+1.5$ 在 $pH 6.0$ 处为 $+0.5$ 在 $pH 9.3$ 处为 -0.5 , 所以组氨酸的 pI 值应是 $pH 6.0$ 和 $pH 9.3$ 的中点, 即 $pI=7.65$ 。

在蛋白质中, 可解离侧链的 pK_a 值可能不同于游离的氨基酸的 pK_a 值。有两个因素可能引起解离常数的变动: 一个因素是 α -氨基和 α -羧基一旦在蛋白质中与其他氨基酸通过肽键连接起来, 就失去了对相邻侧链的诱导效应; 另一个因素是一个可解离侧链在蛋白质的三维结构内的位置可以影响它的 pK_a 值。例如, 核糖核酸酶有 4 个组氨酸残基, 每个侧链的 pK_a 值都稍有不同, 这主要是它们所处的微环境不同造成的。

1.1.3 使用离子交换层析可将各种氨基酸分开

层析(chromatography)也称为色谱(来自 chroma)，其基本原理是分析样品作为流动相流过固相时，样品中的各个成分与固相进行不同程度的相互作用，使得样品中的各个成分在固相中的迁移率产生了差别，从而达到分离样品的目的。

根据固相与样品中各个成分之间相互作用的种类的不同，又有离子交换层析、分子筛层析和吸附层析等各种层析方法。通常都是将固相装到一个柱子里，然后使流动相流过柱子(有时需要加压力)，这样的层析方法叫做柱层析(column chromatography)。

如果被分离的样品中各个成分受溶液 pH 的影响，有时带正电荷，有时带负电荷，此时就可利用离子交换层析方法进行样品的分离。离子交换层析的固相是离子交换剂，其中有离子交换树脂、离子交换纤维素、离子交换葡聚糖。通常是以树脂、纤维素、葡聚糖作为基质。这些基质具有网状结构，将带电基团引入到基质上，就形成了离子交换树脂。根据引入的带电基团不同，又分为阳离子交换剂和阴离子交换剂。引入 $\text{SO}_3^- \text{Na}^+$ 或 $\text{COO}^- \text{H}^+$ 基团的叫做阳离子交换剂；引入 $\text{N}^+ (\text{C}_2\text{H}_5)\text{Cl}^-$ 基团的叫做阴离子交换剂。

分离氨基酸常用的是带有耐酸性非常强的磺酸根 $\text{SO}_3^- \text{Na}^+$ (以盐的形式出现) 的强阳离子交换树脂。首先将这种树脂填充到柱子中，然后注入含有样品的流动相，样品中含有阳离子成分 X^+ ，通过静电吸引，与树脂中的带电基团相互作用，结果 X^+ 与 Na^+ 交换，即发生阳离子交换后，形成 $\text{SO}_3^- \text{X}^+$ (图 1.5)。

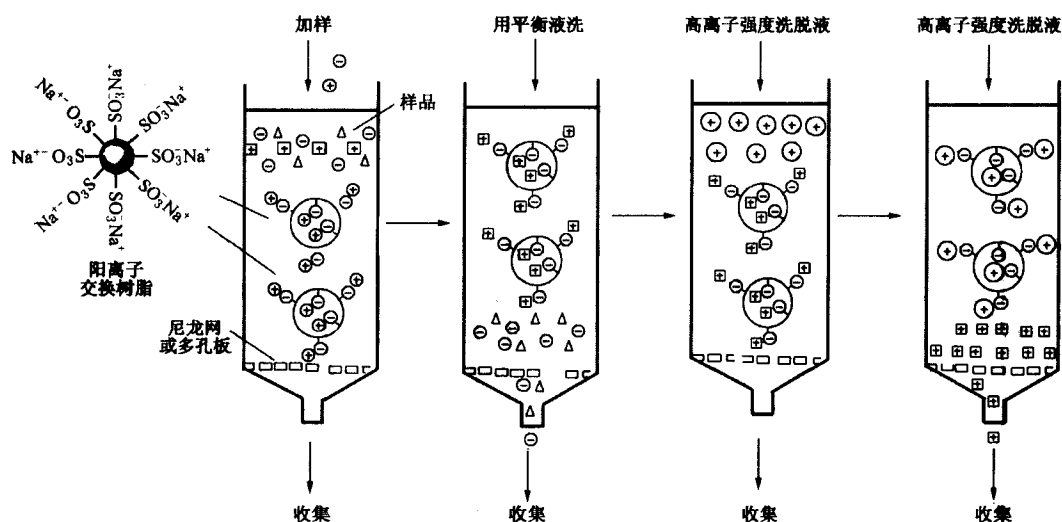


图 1.5 阳离子交换层析示意图

由于发生了离子交换，延缓了溶于流动相的已电离的溶质的移动，同时 X^+ 与 SO_3^- 结合的牢固程度与样品中的 X^+ 成分带的净电荷多少有关，带的净电荷多的成分与树脂的作用强。在样品与树脂充分交换后，即加完样品后，可通过提高流动相中的盐浓度，或是通过改变流动相的 pH，或是同时采用这两种方法，就可以将结合于树脂上的 X^+ 成分，按照它们与树脂结合的强弱程度不同逐一地洗脱下来。图 1.6 给出了氨基酸混合样品的

阳离子交换树脂自动记录层析图。

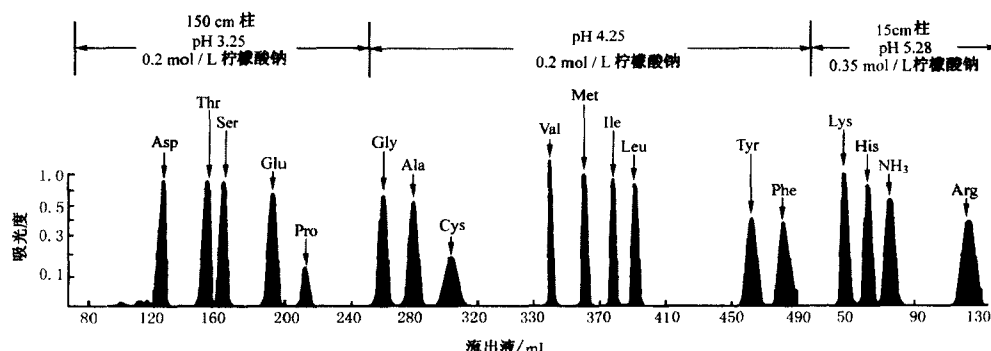


图 1.6 氨基酸混合样品的阳离子交换树脂自动记录层析图

1.1.4 氨基酸可进行特征的化学反应

氨基酸的 α -氨基、 α -羧基以及氨基酸的各种侧链基团可进行很多种化学反应。我们不一一讨论这些化学反应，而着重描述几个生物化学中广泛应用的鉴定和测量氨基酸的重要反应。

氨基酸与茚三酮(ninhydrin)的反应是检测和定量氨基酸和蛋白质的重要反应。茚三酮在弱酸性溶液中与氨基酸共热，所有具有游离氨基的氨基酸都生成紫色化合物 (λ_{470}) (图 1.7)。然而，由于脯氨酸是一个亚氨基酸，所以与茚三酮反应生成的是黄色化合物 (λ_{440})。在一定的反应条件下，产生的颜色的强度(溶液中的光吸收)与氨基酸浓度成比例。所以根据测得的溶液的吸光度，可以算出相应的氨基酸和蛋白质的浓度。

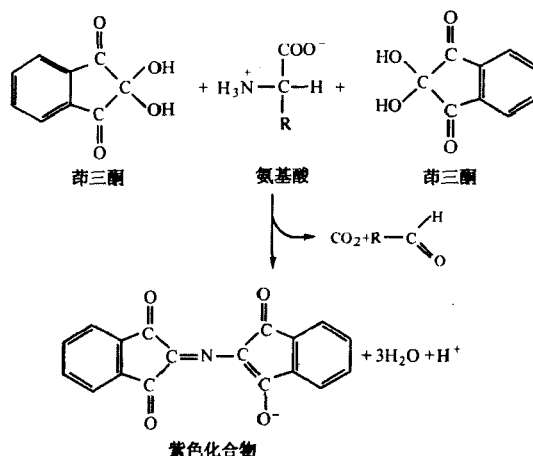


图 1.7 氨基酸和茚三酮的反应

有几种与氨基酸的氨基反应的试剂常用来鉴定蛋白质和多肽的 N-末端氨基酸残基。2,4-二硝基氟苯(2,4-dinitrofluorobenzene, DNFB),也叫做 Sanger 试剂,在弱碱性溶液中与氨基酸发生取代反应,生成黄色化合物二硝基苯基氨基酸(dinitro phenyl amino

acid, DNP 氨基酸)(图 1.8(a))。

丹磺酰氯 (dansyl chloride, DNS-Cl) 是 5-二甲氨基萘-1-磺酰氯 (5-dimethylaminonaphthalene-1-sulfonyl chloride) 的简称。丹磺酰氯与氨基酸反应生成荧光性质强和稳定的丹磺酰衍生物 (DNS-氨基酸)(图 1.8(b))，也常用于多肽链的 N 末端氨基酸的鉴定。

苯异硫氰酸酯 (phenylisothiocyanate, PITC) 在弱碱性条件下，与氨基酸反应生成苯乙内酰硫脲 (phenylthiohydantoin, PTH) 衍生物，即 PTH-氨基酸 (图 1.8(c))，此反应又称为 Edman 反应。该反应是蛋白质或多肽氨基酸序列测定常用的反应。

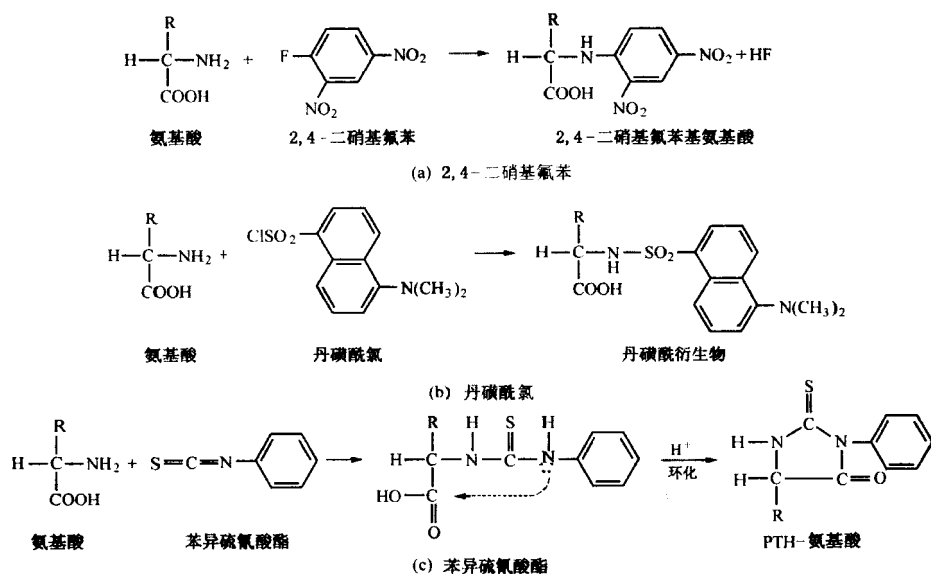


图 1.8 与氨基酸的 α -氨基反应的试剂

1.2 蛋白质中的氨基酸是通过肽键连接的

一个氨基酸的 α -羧基与另一个氨基酸的 α -氨基缩合，通过形成的酰胺键将两个氨基酸连接在一起，这个酰胺键称为肽键 (peptide bond)(图 1.9)，氨基酸缩合生成的产物称为肽 (peptide)。

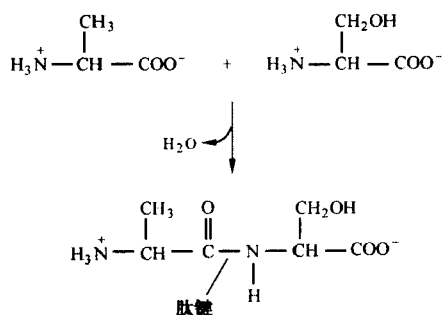


图 1.9 两个氨基酸之间的肽键

蛋白质就是氨基酸通过肽键连接形成的聚合物。在氨基酸缩合的反应中失去了一分子水，相连的氨基酸成分被称为氨基酸残基，两个氨基酸缩合后的产物称为二肽，像图 1.9 那样由丙氨酸和丝氨酸缩合形成的二肽被称为丙氨酰丝氨酸，因为丝氨酸的氨基已被丙氨酰取代了。

在一个肽链主链的两端各有一个游离的氨基和羧基。游离的氨基端称为 N 端 (氨基端)，而游离的羧基端称为 C 端 (羧基端)。在中性 pH 条件下，每一端都带有离子电荷。为方便起

见，肽链中氨基酸残基都是从 N 端至 C 端编号的，所以肽链中氨基酸残基的次序都是从左向右的。

一个蛋白质或多肽的一级结构是通过肽键连接的氨基酸残基的线性序列。氨基酸的 3 个英文字母缩写和一个字母缩写形式常用于表示肽链的一级结构，如一个肽链用氨基酸 3 字母书写，其序列为：Ala-Lys-Gly-Arg-Phe，相应的单字母表示是：AKGRF。二肽、三肽、寡肽和多肽分别指的是由 2 个、3 个、3 个以上和很多（20 个以上）氨基酸残基组成的肽链。二肽含有一个肽键，三肽含有两个肽键，以此类推。一般来说，一条肽链，无论它多长，都有一个游离的 α -氨基和一个游离的 α -羧基，但共价修饰的末端和环形的肽链除外。蛋白质分子的大多数离子电荷都是由它的组成氨基酸残基的侧链贡献的，所以一个蛋白质的离子特性和它的溶解性都取决于它的氨基酸组成。此外，就像我们将在下面看到的那样，氨基酸残基侧链之间的相互作用对于稳定一个蛋白质分子的三维结构有重要贡献。

有些肽比较大，例如胰岛素就是含有 51 个氨基酸残基的多肽，具有重要的生物学活性。有些肽虽然比较小，也具有重要的生理功能，例如加压素（9 肽）和催产素（9 肽）。一些神经多肽的类似物，如内啡肽，就是一种天然的止痛药。还有一些非常简单的肽也常用作食物的调味剂，如甜味剂 Aspartame 就是天冬氨酸苯丙氨酸的甲基酯（图 1.10）。它的甜度是蔗糖的 200 倍，广泛用于食品饮料中。

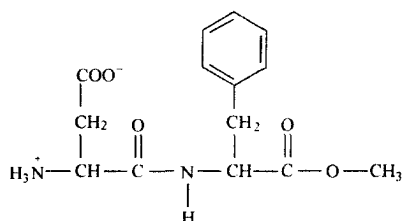


图 1.10 Aspartame 的结构

1.3 肽可以化学合成

多肽合成的基本过程是 20 世纪初由 Emil Fischer 设计的，称为液相合成法，肽链从 N 端向 C 端方向延伸。首先，一个氨基酸的羧基和另一个氨基酸的氨基被保护基团封闭，然后，参与肽键形成的羧基被激活，如形成酰氯，或酸酐。激活的羧基受到游离氨基的亲核攻击形成一个封闭的二肽（图 1.11）。最后，通过水解有选择地除去保护基团而留下一个完整的肽键。

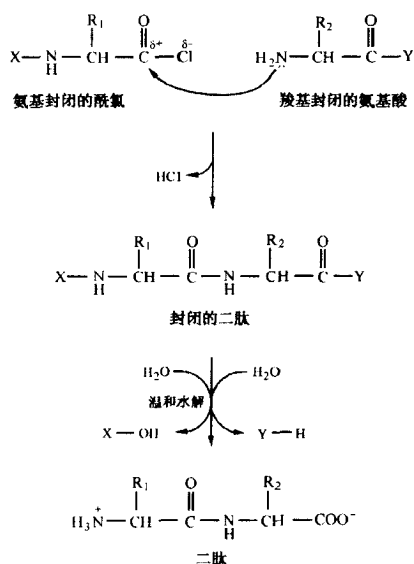


图 1.11 二肽的合成

在肽合成的技术方面取得了突破性进展的是 R. Bruce Merrifield 他设计了一种肽的合成途径并定名为固相合成途径。由于 R. Bruce Merrifield 在肽合成方面的贡献，1984 年获得了诺贝尔奖。图 1.12 给出了肽固相合成途径的简单过程（合成一个二肽的过程）。

氯甲基聚苯乙烯树脂作为不溶性的固相载体，首先将一个氨基被封闭基团（图中的 X）保护

的，首先将一个氨基被封闭基团（图中的 X）保护