

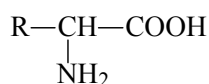
## 第十六章 氨基酸 蛋白质 核酸

蛋白质和核酸都是生物高分子化合物，而氨基酸则是蛋白质的基本组成单位。它们存在于生命细胞中，相辅相成，自动调节，构成了生命活动的物质基础。即使简单的生物如某些病毒，也是由蛋白质和核酸共同组成。细胞内除水分外，其余 80% 的物质为蛋白质。现已证实，生物的遗传特征主要由核酸决定，而生物体的各种生命活动则主要依靠蛋白质来实现。

### § 16.1 氨基酸

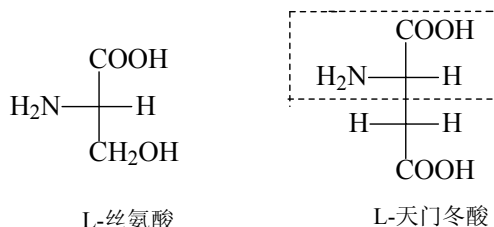
#### 一、氨基酸的结构、分类、名称和物理性质

氨基酸(amino acid)是一类具有特殊重要意义的化合物，因为它们之中许多是与生命起源和生命活动密切相关的蛋白质的基本组成单位。 $\alpha$ -氨基酸是组成所有蛋白质的基本结构单位。天然蛋白质水解的最终产物就是  $\alpha$ -氨基酸（个别除外），所以可以认为  $\alpha$ -氨基酸是蛋白质的基石。组成蛋白质的天然氨基酸有 20 余种，它们在化学结构上都具有共同的特点，就是在羧基邻位  $\alpha$ -碳原子上又一个氨基，因此称之为  $\alpha$ -氨基酸，可以表示为：

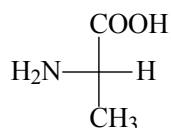


当然，根据氨基在烃基的不同位置，我们可以将氨基酸划分为  $\alpha$ -、 $\beta$ -、 $\gamma$ -、 $\cdots$ 、 $\omega$ -氨基酸。不过最为重要的还是  $\alpha$ -氨基酸。

我们可以定义分子中同时具有氨基和羧基的化合物为氨基酸。除少数氨基酸（如甘氨酸）外它们都是手性分子，因此具有旋光性。它们的相对构型可用 D/L 标记。例如：



并且它们都属于 L 构型。对于大多数氨基酸来说，L 构型就相当于绝对构型(R/S 标记)中的 S 型。例如：

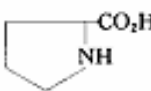
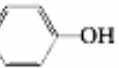
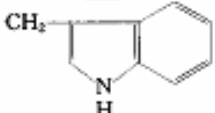
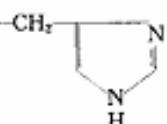


L-丙氨酸 或 S-丙氨酸

在  $\alpha$ -氨基酸中，根据分子中所含的氨基和羧基的数目分为中性、酸性和碱性氨基酸。分子中氨基和羧基数目相等的为中性氨基酸；氨基的数目少于羧基的为酸性氨基酸；氨基的数目多于羧基的为碱性氨基酸。

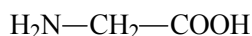
常见的 20 种氨基酸的名称、结构和分类见表 16.1

表 16.1 常见氨基酸

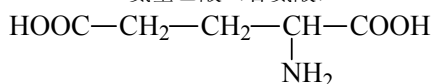
| R 的结构                                                                                                  | 名称                    | 缩写                       | 等电点 (20℃) |
|--------------------------------------------------------------------------------------------------------|-----------------------|--------------------------|-----------|
| <b>中性氨基酸</b>                                                                                           |                       |                          |           |
| -H                                                                                                     | 甘氨酸 glycine           | Gly (甘)                  | 5.97      |
| -CH <sub>3</sub>                                                                                       | 丙氨酸, alanine          | Ala (丙)                  | 6.00      |
| -CH(CH <sub>3</sub> ) <sub>2</sub>                                                                     | * 缬氨酸, valine         | Val (缬)                  | 5.96      |
| -CH <sub>2</sub> CH(CH <sub>3</sub> ) <sub>2</sub>                                                     | * 亮氨酸, leucine        | Leu (亮)                  | 5.98      |
| -CH(CH <sub>3</sub> )CH <sub>2</sub> CH <sub>3</sub>                                                   | * 异亮氨酸, isoleucine    | Ile (异亮)                 | 6.02      |
|  (氨基酸的结构)             | 脯氨酸, proline          | Pro (脯)                  | 6.30      |
| -CH <sub>2</sub> C <sub>6</sub> H <sub>5</sub>                                                         | * 苯丙氨酸, phenylalanine | Phe (苯丙)                 | 5.48      |
| -CH <sub>2</sub> -    | 酪氨酸, tyrosine         | Tyr (酪)                  | 5.66      |
| -CH <sub>2</sub> -    | * 色氨酸, tryptophan     | Trp (色)                  | 5.98      |
| -CH <sub>2</sub> OH                                                                                    | 丝氨酸, serine           | Ser (丝)                  | 5.68      |
| -CH(OH)CH <sub>3</sub>                                                                                 | * 苏氨酸, threonine      | Thr (苏)                  | 5.60      |
| -CH <sub>2</sub> CONH <sub>2</sub>                                                                     | 天冬酰胺, asparagine      | Asn (天冬酰胺)               | 5.41      |
| -CH <sub>2</sub> CH <sub>2</sub> CONH <sub>2</sub>                                                     | 谷氨酰胺, glutamine       | Gln (谷-NH <sub>2</sub> ) | 5.65      |
| -CH <sub>2</sub> SH                                                                                    | 半胱氨酸, cysteine        | Cys (半胱)                 | 5.07      |
| -CH <sub>2</sub> CH <sub>2</sub> SCH <sub>3</sub>                                                      | * 蛋氨酸, methionine     | Met (蛋)                  | 5.74      |
| <b>酸性氨基酸</b>                                                                                           |                       |                          |           |
| -CH <sub>2</sub> CO <sub>2</sub> H                                                                     | 天冬氨酸, aspartic acid   | Asp (天冬)                 | 2.77      |
| -CH <sub>2</sub> CH <sub>2</sub> CO <sub>2</sub> H                                                     | 谷氨酸, glutamic         | Glu (谷)                  | 3.22      |
| <b>碱性氨基酸</b>                                                                                           |                       |                          |           |
| -CH <sub>2</sub> CH <sub>2</sub> CH <sub>2</sub> CH <sub>2</sub> NH <sub>2</sub>                       | * 赖氨酸, lysine         | Lys (赖)                  | 9.74      |
| -CH <sub>2</sub> CH <sub>2</sub> CH <sub>2</sub> NHCNH <sub>2</sub><br> <br>NH                         | 精氨酸, arginine         | Arg (精)                  | 10.76     |
| -CH <sub>2</sub> -  | 组氨酸, histidine        | His (组)                  | 7.59      |

注明: 表中带 \* 号的八种氨基酸在人体内不能合成, 必须从食物中摄取, 所以称为必需氨基酸。

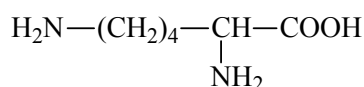
例如: 甘氨酸、谷氨酸、赖氨酸分别是中性、酸性、碱性氨基酸。



氨基乙酸 (甘氨酸)



α-氨基戊二酸 (谷氨酸)



α,ω-二氨基己酸 (赖氨酸)

由于氨基酸分子中同时含有氨基和羧基, 既是酸又是碱, 是两性化合物。在固态时, 它们都是以两性离子或内盐形式存在, 是不易挥发的无色晶体。正是由于氨基酸的内盐性质。

它具有高度的极性，分子间的静电吸引力导致其晶格结构较强，使其熔点较高（一般在 200 °C 以上），并分解生成胺和二氧化碳。一般能溶于水而难溶于苯、乙醚等非极性有机溶剂中。

问题 15.1. 写出下列氨基酸的结构式：

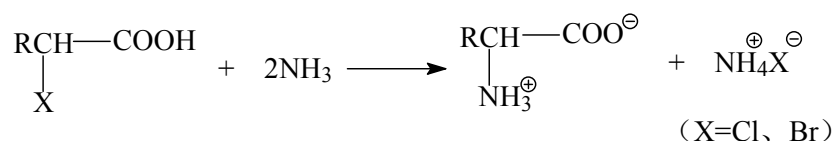
- (1) L-丙氨酸      (2) L-苏氨酸      (3) L-半胱氨酸

问题 15.2. 酸性氨基酸、中性氨基酸、碱性氨基酸的分类原则是什么？它与通常的酸、碱、中性的划分有何差别？

## 二、氨基酸的制备

### 1. α-卤代酸的水解

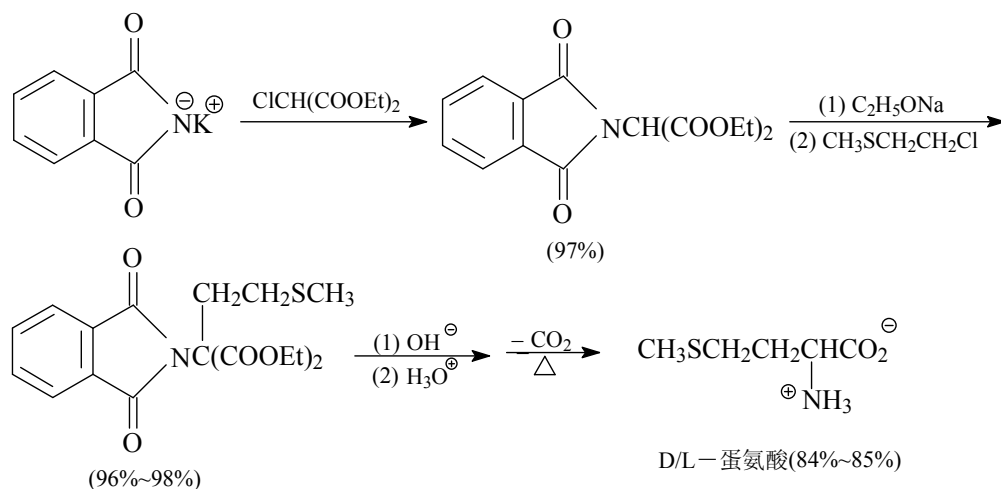
α-卤代酸与过量的氨作用可以得到 α-氨基酸：



α-氨基酸中的氨基的碱性比脂肪胺弱，进一步烷基化的倾向较小，可以得到较纯的 α-氨基酸，但由于产量不高，用得并不多。

### 2. Gabriel 法

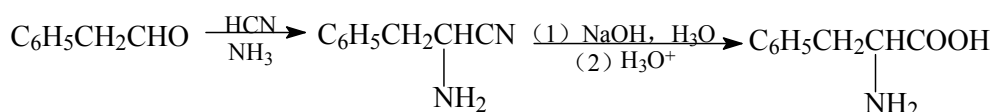
与伯胺的制备类似【见 § 11.5(四)1】，改良的 Gabriel 法也可用于 α-氨基酸的合成。例如卤代丙二酸酯与邻苯二甲酰亚胺的盐作用生成 N-丙二酸酯邻苯二甲酰亚胺，后者经烷基化、水解脱羧可得到 α-氨基酸：



上述合成得到的 α-氨基酸是外消旋体，产物需要拆分才能得到有实际用途的 L-氨基酸。

### 3. Strecker 合成法

醛与氨及氢氰酸反应，可以得到 α-氨基腈，后者水解转化为 α-氨基酸，这种合成方法称为 Strecker 合成法，例如：

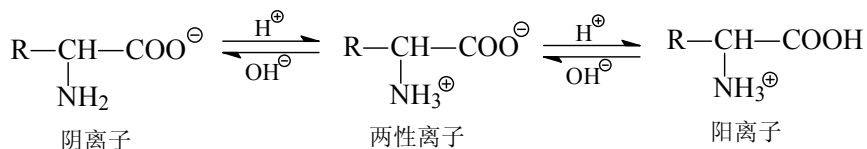


## 三、氨基酸的化学性质

氨基酸分子中既含有氨基又含有羧基，因此它们应该具有胺类和羧酸类的某些典型性质；又由于这两种官能团在分子内的相互影响，因而又表现出一些综合的特性。

### 1. 氨基酸的酸碱性与等电点

氨基酸分子中既含有碱性的氨基，又含有酸性的羧基，所以呈现两性。除本身可以形成内盐外，也可以既能与酸反应生成铵盐，又能与碱反应生成羧酸盐。例如：

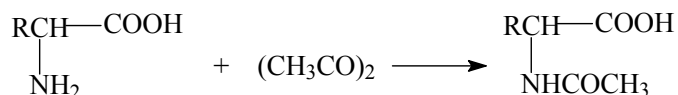


在水溶液中，上面平衡式的移动和溶液的 pH 值有关。在强酸性溶液中，氨基酸以阳离子状态存在，当电解时，它移向阴极；在强碱溶液中，氨基酸以阴离子状态存在，电解时，它移向阳极。若调节电解池的 pH 值，使氨基酸既不向阴极移动也不移向阳极，这个 pH 值称为该氨基酸的等电点 (isoelectric point, IP)。值得注意的是，等电点并不是中性点。不同的氨基酸具有不同的等电点。一般中性氨基酸的等电点在  $\text{pH} \approx 6$ ，酸性氨基酸的等电点  $\text{pH} \approx 3$ ，碱性氨基酸的等电点  $\text{pH} \approx 10$ 。

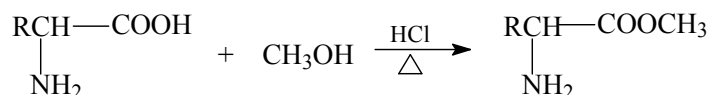
在等电点时，氨基酸主要以两性离子存在，此时氨基酸在水中的溶解度最小。因此用调节等电点的方法，可以从氨基酸的混合物中分离出某些氨基酸。可以用两方法：(1) 利用不同氨基酸的等电点不同，并且在等电点时的溶解度最小，可以将某些氨基酸从混合溶液中沉淀出来；(2) 利用在同一 pH 的不同的氨基酸所带电荷的不同而进行分离。

### 2. 酯化和酰化的反应

$\alpha$ -氨基酸分子中的氨基具有典型氨基的性质，例如可以发生酰基化反应：

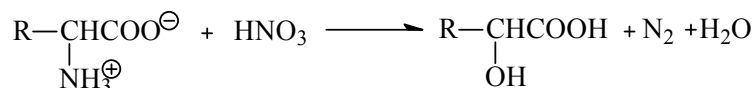


$\alpha$ -氨基酸分子中的羧基也具有羧基所固有的性质，例如可以发生酯化反应：



### 3. 与亚硝酸反应

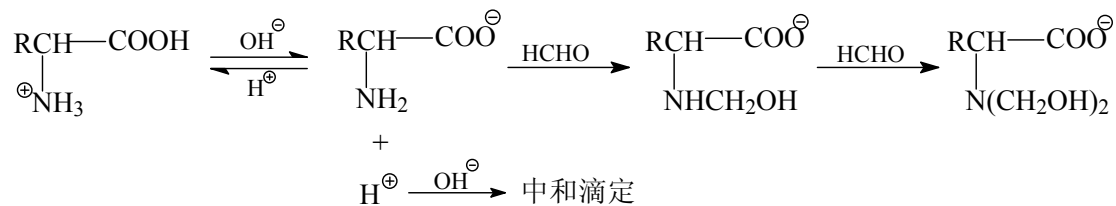
氨基酸与  $\text{HNO}_2$  作用，生成羟氨酸，同时放出  $\text{N}_2$ ，反应迅速且能定量完成。



这样我们测定放出的氮气的体积，就可以计算出氨基酸的含量。这个方法叫做范斯莱克 (Van Slyke) 氨基测定法。

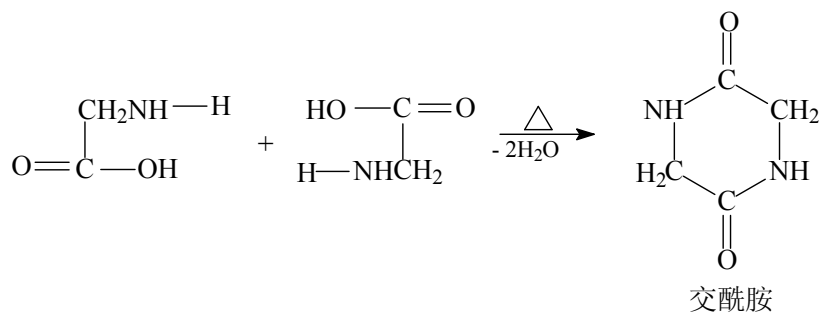
### 3. 与甲醛反应

$\alpha$ -氨基酸的氨基可以与甲醛迅速反应，释放出  $\text{H}^+$ 。然后可以酚酞作指示剂，用  $\text{NaOH}$  滴定间接测定氨基的含量。

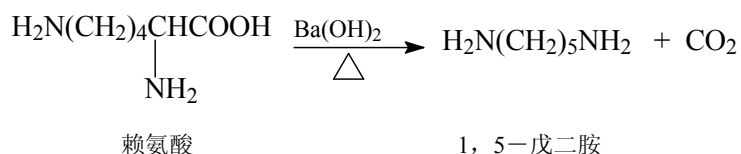


#### 4. 受热后的反应

与  $\alpha$ -羟基酸类似， $\alpha$ -氨基酸受热后生成交酰胺，例如：

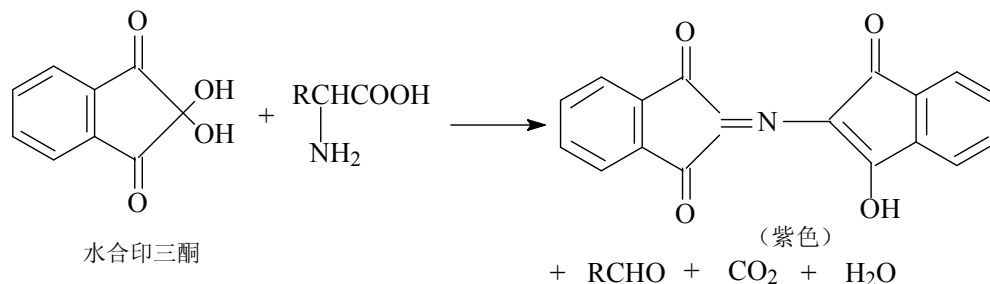


如果在碱性条件下加热，则发生脱羧反应，例如：



#### 5. 显色反应

$\alpha$ -氨基酸的水溶液可以和水合茚三酮反应，生成兰紫色的物质。该颜色反应可以用来鉴别  $\alpha$ -氨基酸（N-取代的  $\alpha$ -氨基酸及  $\beta$  或  $\gamma$ -氨基酸都不发生该颜色反应），也常用作  $\alpha$ -氨基酸的比色测定和层析的显色。在反应中  $\alpha$ -氨基酸被氧化成醛，同时脱去氨基。



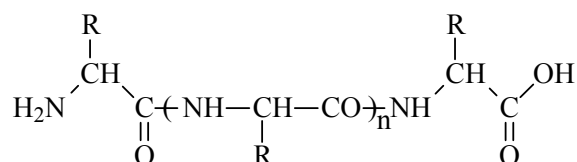
$\alpha$ -氨基酸与水合茚三酮的反应十分灵敏，几微克  $\alpha$ -氨基酸就能显色，所以常用水合茚三酮为显色剂，定性鉴定  $\alpha$ -氨基酸。当然也可以用紫外分光光度法定量测定  $\alpha$ -氨基酸的含量，因为生成的紫色溶液在 570nm 有强吸收，其强度与参与反应的氨基酸的量成正比。

习题 15.3 写出下列反应的主要产物：

- |                                                                                                                        |                                                                                                                                |
|------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------|
| (1) $\text{CH}_3\text{CHCOOC}_2\text{H}_5 + \text{H}_2\text{O} \xrightarrow[\Delta]{\text{HCl}}$<br> <br>$\text{NH}_2$ | (2) $\text{CH}_3\text{CHCOOH} + \text{CH}_3\text{CH}_2\text{COCl} \longrightarrow$<br> <br>$\text{NH}_2$                       |
| (3) $\text{CH}_3\text{CHCONH}_2 + \text{HNO}_2(\text{过量}) \longrightarrow$<br> <br>$\text{NH}_2$                       | (4) 丙氨酸 $\xrightarrow{\Delta}$                                                                                                 |
| (5) $\text{NH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{COOH} \xrightarrow{\Delta}$                                     | (6) $\begin{array}{c} \text{CH}_2\text{COOH} \\   \\ \text{NH}_2 \cdot \text{HCl} \end{array} + \text{SOCl}_2 \longrightarrow$ |

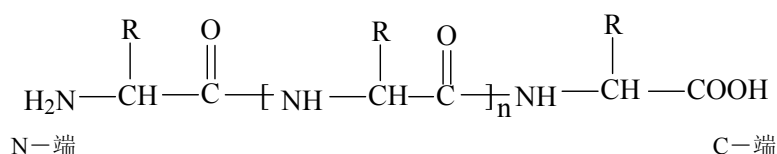
#### 四、肽

氨基酸之间通过肽键相互连接而成的化合物叫做肽。其中由两个氨基酸组成的叫二肽，由三个氨基酸组成的叫三肽，以此类推。一般十肽以下的统称寡肽或低聚肽。而十一肽以上称为多肽，其中分子量超过  $1 \times 10^4$  的多肽称为蛋白质，因此，蛋白质的化学本质就是多肽，用通式可表示为：

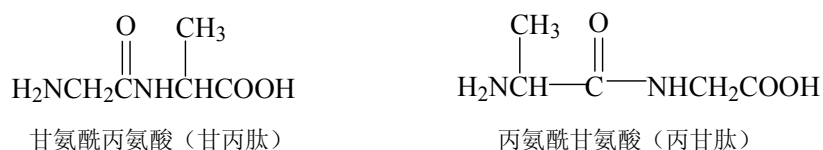


多肽

在肽链中，带有游离氨基的一端称为 N-端，带有游离羧基的一端称为 C-端。例如：

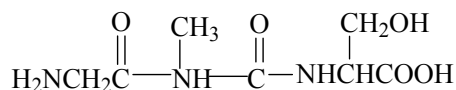


多肽的命名是以 C-端的氨基酸为母体，把肽链中其它氨基酸中的“酸”字改为“酰”字，按在链中的顺序依次写在母体名称的前面。也可用缩写或符号表示，例如：



甘氨酸

丙氨酸



甘氨酸-丙氨酸-丝氨酸

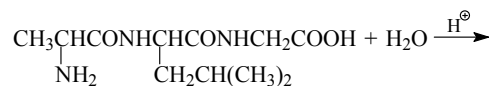
肽中的  $\left(\text{NH}-\underset{\text{R}}{\text{CH}}-\text{CO}\right)$  单位称为氨基酸残基，连接氨基酸残基的酰胺键就是肽

键，氨基酸残基通过肽键形成的链，称为肽链。所以多肽和蛋白质之间没有严格的界限。例如胰岛素，虽然它只含有 51 个氨基酸单位，但一般也认为它是一个小蛋白质。

习题 15.4 由氨基酸组成的多肽其命名原则是什么？

习题 15.5 写出丙氨酸-甘氨酸-丙氨酸的结构式

习题 15.6 写出下列反应的主要产物的结构式：



由于肽键中 C=O 双键中的  $\pi$  电子云与 N 原子上的未共用电子对发生“电子共振”，使肽键具有部分双键的性质，不能自由旋转。如图 16.1 所示。

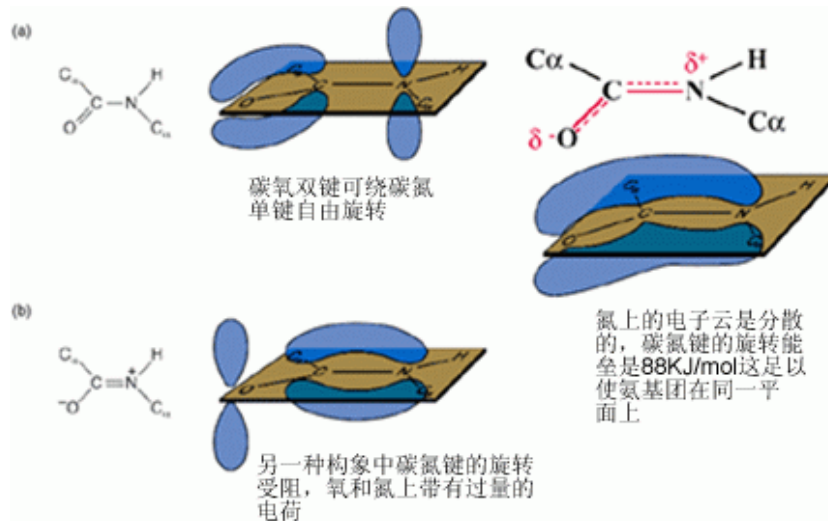


图 16.1 肽键的空间结构

另外与肽键相连的六个原子构成刚性平面结构，称为肽单元或肽键平面。但由于  $\alpha$ -碳原子与其他原子之间均形成单键，因此两相邻的肽键平面可以作相对旋转。此单键的旋转决定两个肽键平面的位置关系，于是肽键平面成为肽链盘曲折叠的基本单位。如图 16.2 所示。

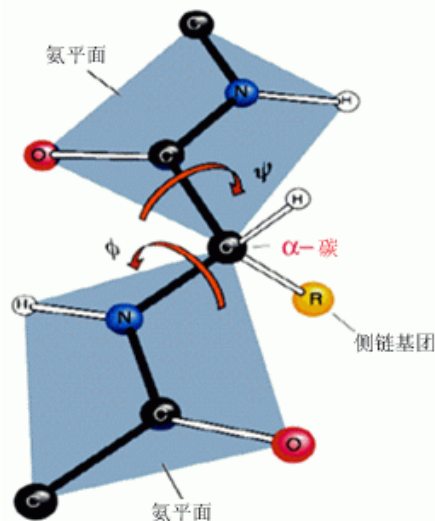


图 16.2 蛋白质立体结构原则

多肽广泛存在于自然界。它们在生物体内起着各种重要的作用，如谷胱甘肽能保护人体，使之免受对细胞有破坏性影响的毒素的侵袭，能延缓人体衰老。因此，半个多世纪以来，有机化学家及生物化学家通力合作，对多肽的研究进行了卓有成效的工作，研究工作主要包括两个方面：一方面需要确定多肽是由哪些氨基酸组成的，各种氨基酸的相对比例是多少，以及它们的排列顺序，然后推测其结构；另一方面要进行合成，以验证所推测的结构是否正确。

和多肽相比，蛋白质具有更长的肽链，并且结构更为复杂。蛋白质部分水解即生成多肽，因此研究多肽结构是了解蛋白质结构的一个重要步骤。

要测定某一个多肽的结构，首先必须要知道这个多肽分子是由哪些氨基酸组成的，然后再确定这些氨基酸在多肽分子中的排列顺序，最后推测其结构。将多肽分子在酸性溶液中进行彻底水解后，便可得到组成多肽的各种氨基酸的混合物，在用色谱法分离并确定它是由哪几种氨基酸组成以及它们的相对含量，然后根据多肽的分子量，就可算出这个多肽中所含的各种氨基酸的数目。至于这些氨基酸在多肽分子中的排列顺序，则是通过端基分析的方法来

确定。关于端基分析法，可以参阅相关书籍。

多肽也是两性离子，也有等电点。在等电点时溶解度最小，多肽的性质与氨基酸极为相似。

## § 16.2 蛋白质

在我们学习了氨基酸之后，现在可以学习蛋白质了(**protein**)。蛋白质在生物体内的功能是多种多样的，如有的负责输送氧气(色蛋白)。有的在新陈代谢中起调节或催化作用(激素或酶)，有的能预防疾病的发生(抗体)，有的则与生物的遗传有关(核蛋白)等。

### 一、蛋白质的元素组成

各种蛋白质经元素分析，发现其元素的成分很近似，测得的百分组成如下： $C=50\% \sim 55\%$ ， $H=6.0\% \sim 7.3\%$ ， $O=19\% \sim 24\%$ ， $N=13\% \sim 19\%$ ， $S=0 \sim 4\%$ ，有的蛋白质中还含有磷、铁、铜、锌、锰等元素。

不同蛋白质中的含氮量相当接近，其平均值为 16%，即 1 克氮相当于 6.25 克蛋白质。这样在分析一个样品的蛋白质含量时，只要测定样品中的含氮量，就可算出其中蛋白质的大致含量。

### 二、蛋白质的性质

蛋白质由氨基酸组成，所以它有与氨基酸类似的性质，比如也有两性性和等电点，在强酸溶液中，以正离子状态存在；在强碱溶液中，以负离子状态存在；在等电点时溶解度最小，可从溶液中析出。蛋白质也可以发生茚三酮显色反应等。此外蛋白质还有一些特殊的性质。

#### 1. 盐析作用

往蛋白质溶液中加入无机盐（如氯化钠、硫酸镁、硫酸铵等），蛋白质便可从溶液中析出，这种作用称为盐析。这是一个可逆过程，盐析出来的蛋白质还可再溶于水，并不影响其性质。所有蛋白质都能发出盐析作用，但不同蛋白质盐析时，所需盐的浓度是不相同的。利用此性质可以分离不同的蛋白质。

用重金属离子  $Hg^{2+}$ 、 $Pb^{2+}$  等形成的不溶性蛋白质，则是不可逆的。

#### 2. 变性

蛋白质在受热、紫外光照射或化学试剂的作用下，性质发生改变，溶解度降低，甚至凝固，这种现象称为蛋白质的变性。变性作用主要是由于蛋白质分子内部结构发生变化所致。例如鸡蛋煮熟就是一个典型的变性过程。硝酸、浓硫酸、三氯乙酸、苦味酸和一些重金属（如  $Hg^{2+}$ 、 $Pb^{2+}$ 、 $Ag^+$  等）都可使蛋白质变性。蛋白质一旦变性，就丧失了其原有的生理功能。

### 三、蛋白质的结构

蛋白质不仅相对分子质量很大，而且结构非常复杂。它不仅存在着多肽链内氨基酸的种类和排列问题(构造)，还存在着一条多肽链本身或几条多肽链之间的空间结构问题(构型和构象)。通常将蛋白质的结构层次分为四级。

#### 1. 一级结构

氨基酸在蛋白质肽链中排列的顺序叫做它的一级结构。肽键是一级结构中的主要连接键。各种蛋白质的生物活性，首先是由它的一级结构决定的。一级结构也是蛋白质最基本的



结构。它是由基因上遗传密码的排列顺序所决定的，各种氨基酸按遗传密码的顺序通过肽键连接起来。每一种蛋白质分子都有自己特有的氨基酸的组成和排列顺序即一级结构，由这种氨基酸排列顺序决定它的特定的空间结构，也就是蛋白质的一级结构决定了蛋白质的二级三级等高级结构，这就是荣获诺贝尔奖的著名的 Anfinsen 原理。图 16.3 给出了胰岛素的一级结构。

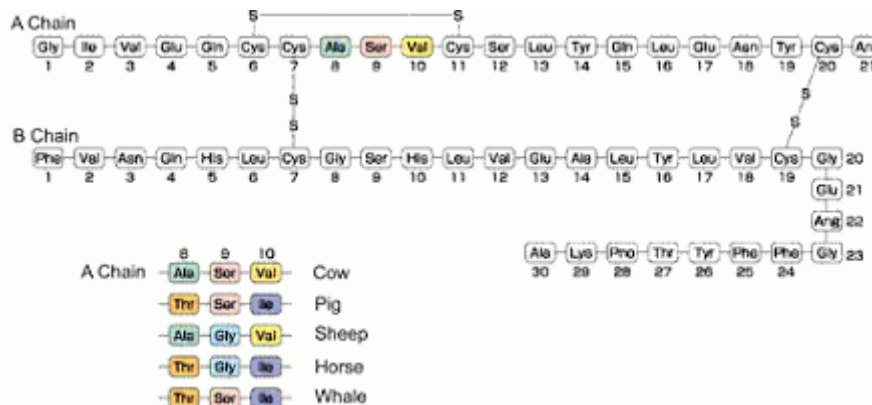


图 16.3 胰岛素的一级结构及不同动物胰岛素在 A 链中的差异

胰岛素 (Insulin) 由 51 个氨基酸残基组成，分为 A、B 两条链。A 链 21 个氨基酸残基，B 链 30 个氨基酸残基。A、B 两条链之间通过两个二硫键联结在一起，A 链另有一个链内二硫键。

## 2. 二级结构

蛋白质分子中的肽链并非是直线型的，而是排列成盘曲或折叠状，这就是蛋白质的二级结构。二级结构主要有两种形式，一种是  $\alpha$ -螺旋形，另一种  $\beta$ -折叠形。在二级结构中有氢键参与，以维持其稳定性。

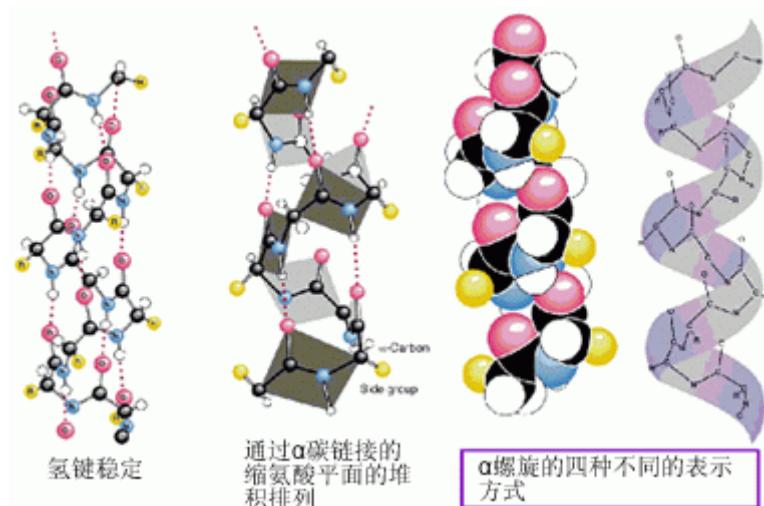


图 16.4 四种不同的  $\alpha$ -螺旋

$\alpha$ -螺旋( $\alpha$ -helix)是蛋白质中最常见最典型含量最丰富的二级结构元件。在  $\alpha$  螺旋中，每个螺旋周期包含 3-6 个氨基酸残基，残基侧链伸向外侧，同一肽链上的每个残基的酰胺氢原子和位于它后面的第 4 个残基上的羰基氧原子之间形成氢键。这种氢键大致与螺旋轴平行。一条多肽链呈  $\alpha$ -螺旋构象的推动力就是所有肽键上的酰胺氢和羰基氧之间形成的链内氢键。在水环境中，肽键上的酰胺氢和羰基氧既能形成内部( $\alpha$ -螺旋内)的氢键，也能与水分子形成氢键。如果后者发生，多肽链呈现类似变性蛋白质那样的伸展构象。疏水环境对于

氢键的形成没有影响,因此,更可能促进 $\alpha$ -螺旋结构的形成。在图 16.4 中我们可以看到 4 种不同的 $\alpha$ -螺旋。

$\beta$ -折叠( $\beta$ -sheet)也是一种重复性的结构,可分为平行式和反平行式两种类型,它们是通过肽链间或肽段间的氢键维系。

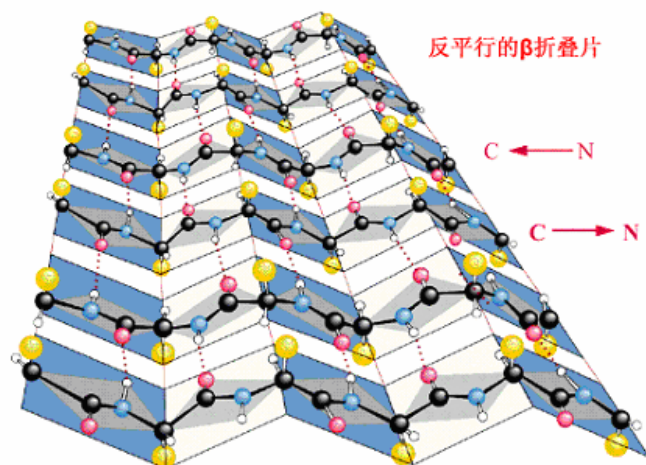


图 16.5  $\beta$ -折叠

可以把它们想象为由折叠的条状纸片侧向并排而成,每条纸片可看成是一条肽链,称为 $\beta$ 折叠股或 $\beta$ 股( $\beta$ -strand),肽主链沿纸条形成锯齿状,处于最伸展的构象,氢键主要在股间而不是股内。 $\alpha$ -碳原子位于折叠线上,由于其四面体性质,连续的酰氨平面排列成折叠形式。需要注意的是在折叠片上的侧链都垂直于折叠片的平面,并交替的从平面上下二侧伸出。平行折叠片比反平行折叠片更规则且一般是大结构而反平行折叠片可以少到仅由两个 $\beta$ 股组成。如图 16.5 所示。

### 3. 三级结构

蛋白质的三级结构是肽链(二级结构)进一步扭曲折叠形成的复杂空间结构。主要针对球状蛋白质而言的是指整条多肽链由二级结构元件构建成的总三维结构,包括一级结构中相距远的肽段之间的几何相互关系,骨架和侧链在内的所有原子的空间排列。在球状蛋白质中,侧链基团的定位是根据它们的极性安排的。蛋白质特定的空间构象是由氢键、离子键、偶极与偶极间的相互作用、疏水作用等作用力维持的,疏水作用是主要的作用力。有些蛋白质还涉及到二硫键。如果蛋白质分子仅由一条多肽链组成,三级结构就是它的最高结构层次。图 16.6 是磷酸丙糖异构酶和丙酮酸激酶的三级结构。

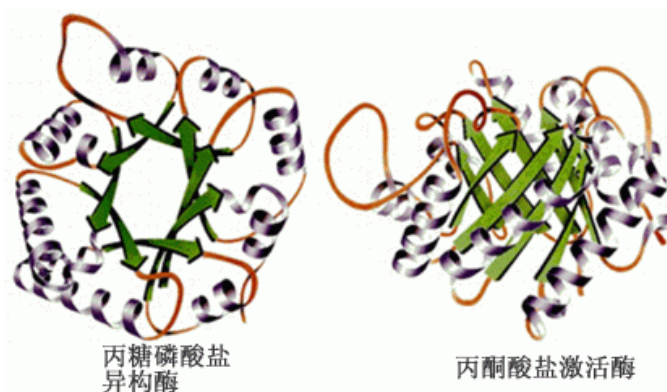


图 16.6 磷酸丙糖异构酶和丙酮酸激酶的三级结构

#### 4. 四级结构

许多球状蛋白质是由一条或多条肽链构成的，每条肽链都有各自一、二、三级结构。这些肽链称为蛋白质的亚基或原体。各个亚基聚合成大分子的方式称为蛋白质的四级结构。稳定四级结构的作用力与稳定三级结构的没有本质区别。亚基的二聚作用伴随着有利的相互作用包括范德华力，氢键，离子键和疏水作用还有亚基间的二硫键。亚基缔合的驱动力主要是疏水作用，因亚基间紧密接触的界面存在极性相互作用和疏水作用，相互作用的表面具有极性基团和疏水基团的互补排列；而亚基缔合的专一性则由相互作用的表面上的极性基团之间的氢键和离子键提供。图 16.7 给出了蛋白质的四级结构。

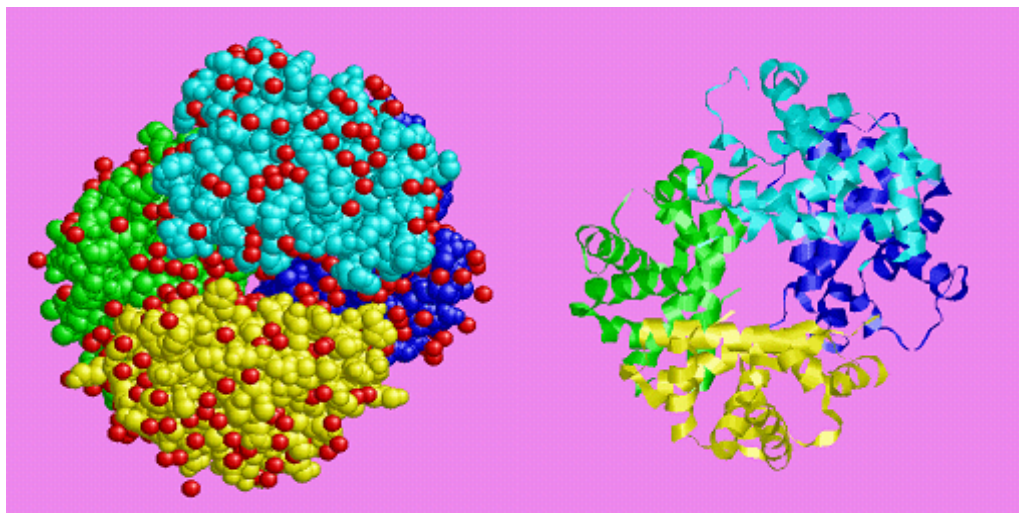


图 16.7 蛋白质的四级结构

#### 四、蛋白质的新来源及其应用前景

人类的食物中，蛋白质是最重要的营养成分，人体的各种生理活动大多数是由蛋白质来承担的。自古以来，蛋白质的主要来源是动物性食物，即肉、鱼、蛋、奶等。随着世界人口的增长，蛋白质供应不足越来越严重。严峻的事实迫使人们去探索生产蛋白质的新途径。科学家们提出三种设想：一是人工合成食品；二是人工模拟生物体内的转化过程来生产食品；三是人工培养微生物来生产蛋白质。且前微生物蛋白的生产已取很可喜的成功。

利用微生物生产蛋白质的原料来源丰富。如石油和煤炭、天然气和淀粉等，还有许多工业废料和一些农副产品的下脚料，甚至城市的垃圾。生产的关键是筛选出有效的酵母菌。高效的微生物繁殖速度极快，如 500kg 的酵母菌在 24h 内能生产出 2 吨多活物质。近年来发现一种叫做氢细菌的微生物，能直接利用  $H_2$ 、 $O_2$  和其它一些简单的无机物来生产蛋白质。这样看来，实现用空气和水来生产粮食不是科学幻想。

生物大分子是分子世界中的一个显赫的大“家族”：相对分子质量大，形状奇特，结构复杂，种类繁多，又各具独特功能。生物大分子主要分为蛋白质、脂类和多糖类，它们是组成生物体的基本元件。如今，对生物大分子感兴趣的除了最终了解生命的本质这个远期的目标外，近期的蛋白质应用方向是由蛋白质构成集成电路，其大小仅为现在的硅集成电路的十万分之一，而开关速度可达一千亿分之一秒。科学家们认为，生物界存在的蛋白质可能有 100 亿种，这是制造生物芯片的极好材料，而生物芯片的使用会给计算机带来一场革命。

## § 16.3 核酸

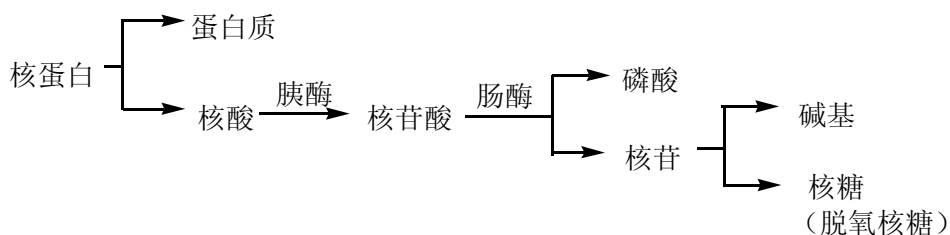
除前面讨论过的蛋白质之外,核酸(Nucleic acid)是对生命现象非常重要的生物大分子。生命活动主要通过蛋白质来体现,但生物体的遗传特征则主要由核酸决定。核酸主要是以与蛋白质结合成核蛋白的形式存在于细胞中,它是1868年瑞士科学家米舍尔(Miescher J F)首先从细胞核中分离出来的具有酸性的物质,所以称为核酸。现在知道,核酸不仅存在于细胞核内、在细胞质,特别是细胞质的粒质中,也含有丰富的核酸。核酸是蛋白质生物合成不可缺少的物质,又是生物遗传的物质基础。可以说,核酸是支配人体整个生命活动的本源物质,被现代科学称之为“生命之源”、“生命之本”。所以是近年来生物化学、有机化学及医学中研究得最广泛、最活跃的课题之一。至今已有38位科学家因涉及核酸的研究而荣获诺贝尔奖。这里仅就有机化学的角度对核酸的组成及两类重要的核酸加以简介。

### 一、核酸的元素组成

经过元素分析,核酸分子中除含C、H、O、N四种元素之外,还含有大量的P,个别的还含有S。其中含N 15%~16%,含P 9%~10%。

### 二、核酸的构成

核蛋白是一种结合蛋白,核酸就是它的非蛋白部分,可由核蛋白水解获得,后者还可以再彻底水解得到碱基和核糖。



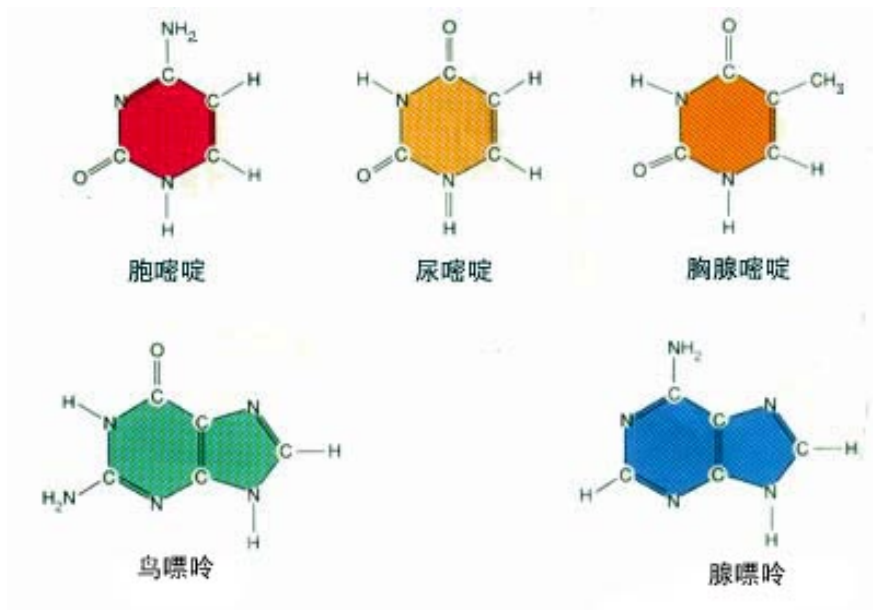
从上可看到核酸是多聚核苷酸,而核苷酸是它的基本结构单位。所以,核酸是由核苷酸组成,而核苷酸又是由碱基、戊糖及磷酸组成。

#### 1. 核糖和脱氧核糖

核酸中的戊糖部分分为核糖及脱氧核糖,所以核酸也分为核糖核酸(ribonucleic acid, 简称为RNA)和脱氧核糖核酸(deoxyribonucleic acid, 简称为DNA)两大类。

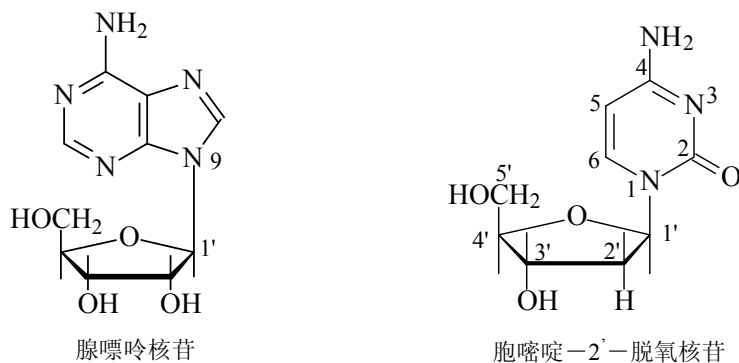
#### 2. 碱基

碱基(base): 构成核苷酸的碱基分为嘌呤(purine)和嘧啶(pyrimidine)二类。前者主要指腺嘌呤(adenine, A)和鸟嘌呤(guanine, G), DNA和RNA中均含有这二种碱基。后者主要指胞嘧啶(cytosine, C)胸腺嘧啶(thymine, T)和尿嘧啶(uracil, U), 胞嘧啶存在于DNA和RNA中, 胸腺嘧啶只存在于DNA中, 尿嘧啶则只存在于RNA中。这五种碱基的结构如下所示:



### 3. 核苷

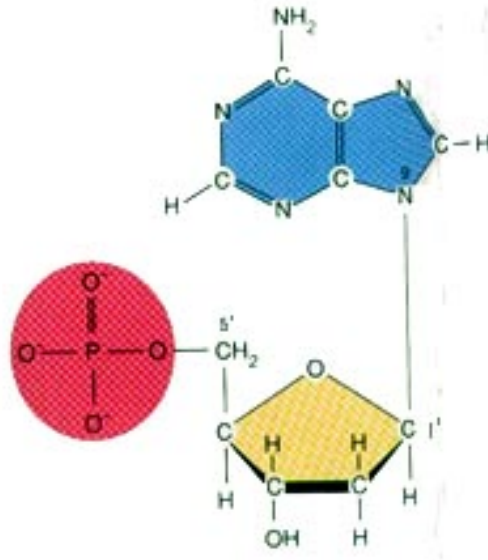
核苷 (nucleoside) 是一种糖苷，由戊糖和碱基缩合而成。戊糖与碱基之间以糖苷键相连接。戊糖分子的 1' 位 C 原子上的羟基与嘧啶碱的 1 位 N 原子或与嘌呤碱的 9 位 N 原子上的氢缩去水后相连接。所以，戊糖与碱基间的连接是 N-C 键，一般称之为 N-糖苷键。用 x 射线衍射法已证明，核苷中的碱基与糖环平面互相垂直。核酸中的主要核苷有八种。如：腺嘌呤核苷。



根据核苷中所含戊糖的不同，将核苷分成两大类：核糖核苷和脱氧核糖核苷。对核苷命名时，必须先冠以碱基的名称。例如腺嘌呤核苷、胞嘧啶-2'-脱氧核苷等。

### 4. 核苷酸

核苷酸是核苷与磷酸残基构成的化合物，即核苷的磷酸酯。核苷酸是核酸分子的结构单元。核酸分子中的磷酸酯键是在戊糖 C-3' 和 C-5' 所连的羟基上形成的，故构成核酸的核苷酸可视为 3'-核苷酸或 5'-核苷酸。DNA 分子中是含有 A,G,C,T 四种碱基的脱氧核苷酸；RNA 分子中则是含 A,G,C,U 四种碱基的核苷酸。下面是 5'-磷酸腺苷的结构式：



### 三、DNA 与 RNA

核酸的相对分子质量因来源不同而异。RNA 主要存在于细胞质中，相对分子质量一般在 20,000~500,000；DNA 主要存在于细胞核中，相对分子质量一般在 6,000,000~300,000,000。

DNA 是由四种主要的碱基：腺嘌呤、鸟嘌呤、胞嘧啶及胸腺嘧啶；而 RNA 是由腺嘌呤、鸟嘌呤、胞嘧啶及尿嘧啶组成。从组成上不难看出，除戊糖不同外，就是 RNA 中没有胸腺嘧啶，而 DNA 中没有尿嘧啶。

核酸是由许多核苷酸以一定顺序连接，并具有特征的三维结构的高分子化合物。其结构复杂，人类尚未完全了解清楚，在此只介绍一级结构和二级结构。

#### 1. 一级结构

DNA 的一级结构是指 DNA 分子中核苷酸的排列顺序，DNA 顺序（或序列）是这一概念的简称。由于核苷酸之间的差异仅仅是碱基的不同，故可称为碱基顺序。DNA 是巨大的生物高分子，如人的 DNA 就包含了  $3 \times 10^9$  碱基对，如此数目的碱基所能容纳的信息量之大是可想而知的。生物世界里形形色色的遗传信息都包含在组成 DNA 的 A,G,C,T 这四种核苷酸的排列顺序之中。DNA 分子中不同排列顺序的 DNA 区段构成特定的功能单位，这就是基因，不同基因的功能各异，各自分布在 DNA 的一定区域。基因的功能取决于 DNA 的一级结构，要想解释基因的生物学含义，就必须弄清 DNA 顺序。因此，DNA 顺序测定是分子遗传学中一项既重要又基本的课题。图 16.8 是 DNA 片段结构示意图。

#### 2. 二级结构

1953 年，华特生 (Watson) 和克里克 (Crick) 以非凡的洞察力，在前人研究的基础上，以立体化学上的最适构型建立了一个与 DNA X 射线衍射资料相符的分子模型 DNA 双螺旋结构模型 (图 16.9)。这是一个能够在分子水平上阐述遗传 (基因复制) 的基本特征的 DNA 二级结构。它使长期以来神秘的基因成为了真实的分子实体，是分子遗传学诞生的标志，并且开拓了分子生物学发展的未来。由双螺旋结构可知，当一条多核苷酸的碱基序列确定后，即可推知另一条主链的碱基序列，这就决定了 DNA 能控制遗传信息，具有高度的保真性。DNA 双螺旋结构的确立，不仅揭示了 DNA 分子结构特征，而且科学地阐明了 DNA 的遗传功能，是现代分子生物学的里程碑。图 16.9 是 DNA 的双螺旋结构。

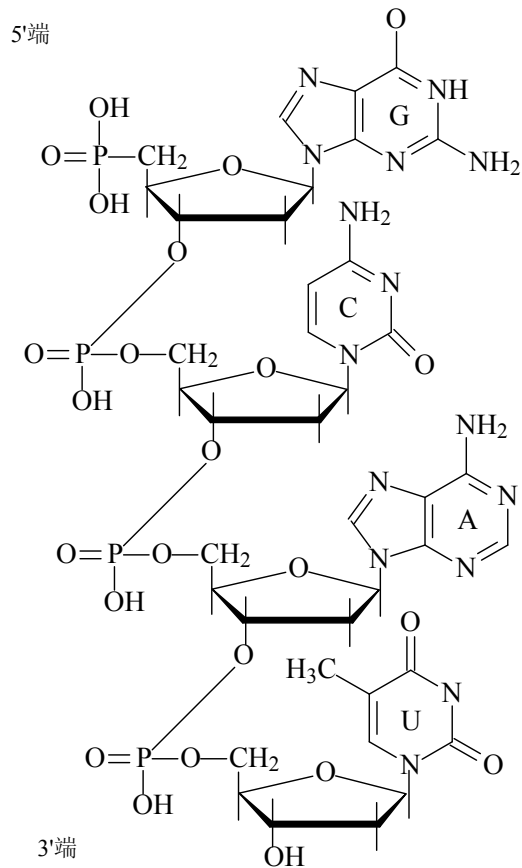


图 16.8 DNA 片段结构示意图

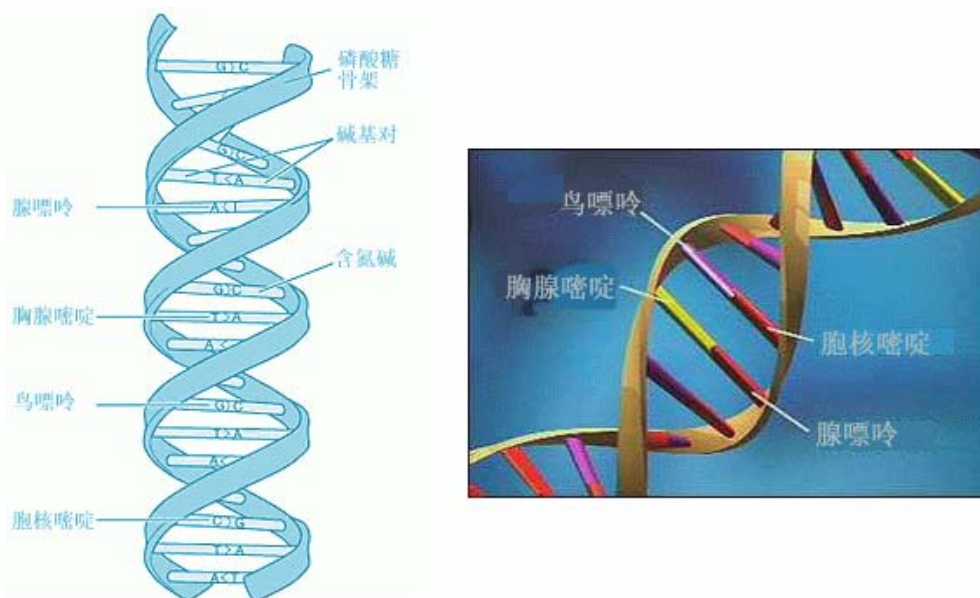


图 16.9 华特生、克里克提出的双螺旋结

习题 15.7 DNA 和 RNA 在结构上有什么主要差别？

习题 15.8 已知甲、乙、丙三种蛋白质的等电点依次是 2.0、7.5、9.0。问：

(1) 它们分别在水溶液中各带何种净电荷？

(2) 若要使这三种蛋白质在混合液中净电荷都为正，你将怎样调节溶液的 pH？该 pH 值至少小于或大于何 pH 值？

## 四、DNA 测序

### 1. 常规 DNA 测序的原理

制作物理图谱的过程是一个逐步精细的过程。第一步把每条染色体分成平均长度在 400kb (kilobase 千碱基对, DNA 片段的长度单位, 一个千碱基对相当于 1000 个核苷) 的长片段, 每段克隆到一个 YAC (酵母人工染色体) 上, 所有 YAC 克隆都按照其在染色体上的实际位置进行排序, 我们就得到了一个能够覆盖整个染色体的 YAC 文库。

把每一个 YAC 克隆携带的染色体片段经部分酶切形成一系列有重叠区域的 40kb 左右的片段克隆到粘粒上, 得到粘粒文库。每个粘粒上的染色体片段再经酶切形成 4kb 左右的片段克隆到测序专用的质粒载体上。测序质粒上携带的 4kb 的片段就可以用现在常规测序的方法进行测序了。把所有质粒克隆的 DNA 片段序列读出, 再按照各个片段在染色体上的实际位置进行排列, 最后就可以得到染色体的全部核苷酸碱基对序列。染色体的 DNA 碱基序列是基因组物理图谱的最精细形式。

所谓“常规测序方法”的基本特点有两个: 第一, 把待测序的 DNA 分子进行处理, 得到每个只差 1 个核苷酸的一系列逐步缩短的 DNA 分子的混合物; 第二, 通过凝胶电泳把这些 DNA 分子分离开来, 形成阶梯状排列的条带, 然后逐个读出 DNA 的碱基序列。

### 2. 化学法测序

得到长度只差一个碱基的 DNA 分子的方法主要有两种。一种是用化学方法把待测序的 DNA 片段在每个碱基处切断一次。这是由 Maxam 和 Gilbert 发明的方法。具体做法是把待测序的 DNA 分子成单链分子, 其 5' 端用  $^{32}\text{P}$  进行放射性标记。然后把这些单链 DNA 分子分成 4 份, 每份用一种化学试剂处理 DNA 片段, 每种试剂可使 DNA 分子在一种碱基的 5' 端的磷酸二酯键处发生断裂。例如, 试剂一可使单链 DNA 分子在 A 碱基处断裂, 试剂二可使单链 DNA 分子在 T 碱基处断裂, 依此类推。把反应条件控制好, 使每个 DNA 分子只发生一次断裂, 这样, 我们就得到 4 种反应产物, 每种由在一种碱基处发生断裂形成的 DNA 片段组成。

把这 4 种反应产物用聚丙烯凝胶电泳进行分离, 两个 DNA 片段只要相差 1 个碱基, 就可以在这种凝胶中被分成两个条带。电泳完成后, 用 X 光胶片进行曝光, 最后得到一张由不同条带组成的序列图。从这张图上就可以读出待测 DNA 片段的碱基序列。5' 端的第一个碱基 G 读不出来, 可以通过测定互补链的序列测出这个碱基。

### 3. 酸法测序与测序的自动化

另外一种得到长度只差一个碱基的 DNA 分子的方法是英国科学家桑格发明的, 这种方法利用 DNA 聚合酶以待测序的 DNA 单链分子为模板合成互补的新链。在合成新链时, 合成原料除了 4 种脱氧核糖核苷酸外还加入一种 2' 和 3' 位上的羟基都脱除的核苷酸。由于缺少 3' 羟基, 当这种核苷酸被结合到链上后, 它的后面不能再结合其他核苷酸, 链的合成就此终止。与化学法测序类似, 我们可以准备 4 种反应物, 加入的核苷酸类似物分别携带 A, G, C, T 碱基, 每种反应物里包含在一种碱基处终止链延伸的长短不同的 DNA 片段。这些 DNA 片段也要用放射性标记, 经过凝胶电泳和放射自显影, 得到 DNA 条带图谱, 根据图谱可以读出 DNA 的碱基序列。

这种方法比化学法简单, 条件易于控制。用 4 种不同的荧光化合物分别标记 4 种反应的产物, 就可以做到把 4 种反应物混合在一起进行电泳, 可以提高电泳分析的效率。这种方法利用现代精密仪器和机器人技术可以实现 DNA 测序的高度自动化。目前市场上已经有各种



型号的 DNA 自动测序仪可供选购。

到 2003 年人类基因组的计划已经完成。这个计划包含了 90% 的人类基因组的序列，每个区域测定 5 次左右。这个序列可以作为“参考基因组”或者“标准基因组”用于生物医学研究。

#### 4. 单分子荧光测序

单分子荧光测序是一种快速 DNA 测序法，这是利用单分子操作技术直接读取 DNA 的碱基序列的方法，与传统的荧光测序法相比，这种方法可大大提高速度。用桑格测序方法现在每天可以解读上万个碱基序列，但是如果单分子荧光测序取得成功，它可以在两分钟内完成传统方法一天的工作。

单分子荧光测序的主要过程如下。第一步，取一条大约有 5 万个碱基那么长的单链 DNA 分子，把它的一端用化学方法连接在一个非常微小的塑料球上，DNA 分子就会缠绕在塑料球上。第二步，在一张类似激光唱片的圆盘上铺一层很薄的液体薄膜，然后用激光光钳把塑料球放在这张光盘的液体膜上进行移动，DNA 分子就会在后面被拖着展开，就好像一只船拖着一条绳子快速前进，把绳子拉展一样。第三步，让拉展的 DNA 分子与一种核酸外切酶结合。这种核酸外切酶可以和 DNA 的游离的末端结合，然后逐个把 DNA 的碱基切割下来。第四步，用单分子光谱技术逐个识别并且读出碱基即达到了测序的目的。

目前，单分子荧光测序技术还没有完全成熟。主要问题是用于识别单个碱基的单分子光谱技术还没有过关。利用隧道扫描显微镜和原子力显微镜直接读取 DNA 分子碱基序列的研究也正在进行之中，近期内有可能取得突破。

#### 5. DNA 芯片与杂交测序

DNA 芯片是一种通过杂交测定未知 DNA 序列的新技术。在一个玻璃或硅片上合成大量的寡聚核苷酸片段，例如可以合成 8 个碱基长的全部可能的寡聚核苷酸片段 ( $4^8 = 65, 536$  种)。这些探针一头固定在固体基质上，另外一端是游离的。它们在硅片上有规律地排列着，每个特定位置上探针的序列都是已知的。

假如有一个 DNA 片段需要测序，我们可以把它的单链形式用荧光进行标记，然后与硅片上的 6 万多种探针进行分子杂交，在荧光显微镜下观察杂交结果。如果某个探针与待测 DNA 的某个部分的序列是完全互补的，待测 DNA 分子就被结合到硅片上，这个探针所在的位置就会发出荧光。这种包含大量的生物遗传信息的寡聚核苷酸阵列就叫 DNA 芯片，也叫基因芯片。

基因芯片利用固定探针与样品进行分子杂交产生的杂交图谱而排列出待测样品的序列，这种测定方法快速而具有十分诱人的前景。Mark chee 等用含 135000 个寡核苷酸探针的阵列测定了全长为 16.6kb 的人线粒体基因组序列，准确率达 99%(7)。Hacia 等用含有 48000 个寡核苷酸的高密度微阵列分析了黑猩猩和人 BRCA1 基因序列差异，结果发现在外显子 11 约 3.4kb 长度范围内的核酸序列同源性在 98.2%到 83.5%之间，提示了二者在进化上的高度相似性。

DNA 芯片的制备要利用微电子芯片生产中的光刻技术以及在位组合化学技术。DNA 芯片的阅读要使用显微术，信息的解读要利用计算机技术。因此，DNA 芯片是多学科交叉的产物。

## 习 题

1. 命名或写出下列化合物的结构式：

(1) L-丙氨酸

(2) L-苏氨酸

(3) L-半胱氨酸

(6) 甘氨酸-亮氨酸 (7) 丙酰胺三肽

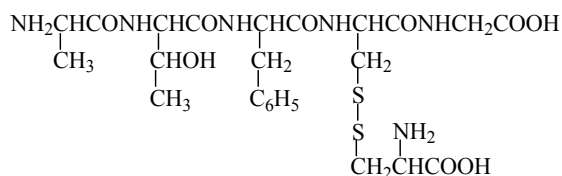
(8)  $\text{CH}_3\text{CH}_2\text{CH}(\text{CH}_3)\text{CH}(\text{NH}_2)\text{COOH}$

(9)  $\text{HOOCCH}_2\text{CH}_2\text{CH}(\text{NH}_2)\text{COOH}$

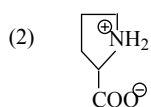
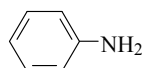
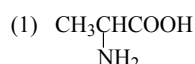
2. 写出下列氨基酸按如下次序相结合所形成的多肽的结构式。

丙·缬·苯丙·甘·亮

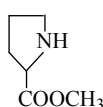
3. 下列多肽水解后, 生成什么样的化合物?



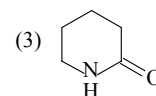
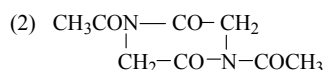
4. 试用简便的方法区别下列各组化合物:



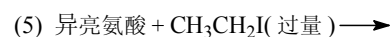
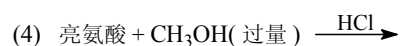
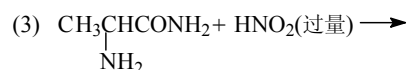
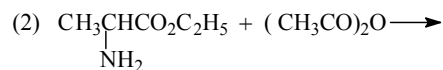
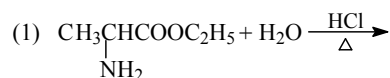
和



5. 下列化合物水解后, 生成哪些化合物?



6. 写出下列反应的主要产物:



7. A 化合物分子式为  $\text{C}_3\text{H}_7\text{O}_2\text{N}$ , 有旋光性, 能与  $\text{NaOH}$  和  $\text{HCl}$  生成盐, 并能与醇生成酯, 与亚硝酸作用是放出  $\text{N}_2$ , 写出 A 化合物的构造式。

8. 化合物 A, 分子式  $\text{C}_5\text{H}_{10}\text{O}_3\text{N}_2$ , 具有旋光性, 它既能与酸反应又能与碱反应, 且能与茚三酮水溶液反应。A 与  $\text{HNO}_2$  反应放出  $\text{N}_2$ , 并得到化合物 B, B 经水解得到化合物 C, 及甘氨酸。C 可发生碘仿反应, 并生成化合物 D, D 可被  $\text{KMnO}_4$  酸性溶液氧化生成二氧化碳和水。

(1) 写出 A、B、C、D 的结构式及反应式。

(2) 写出 A 的两种构型式, 并用 R、S (或 D、L) 标出其构型。