



# DNA 及其结构

董贻诚

现在人们都知道核酸包括核糖核酸 (RNA) 和脱氧核糖核酸 (DNA) 两大类, 而 DNA 是细胞内最主要的遗传物质, 然而这一科学知识的获得却经过了长期的曲折过程。核酸的发现最初是从细胞核的组成研究开始的。1868 年瑞士青年科学家 F. Miescher 在德国 Tübingen 从外科医院收集绷带作为材料, 从绷带上的浓细胞的细胞核中抽提出一类含磷元素较丰富的酸性化合物, 称它为“核质”。这实际上已是含有蛋白质的核酸制品了。20 年后 R. Altmann 制得了不含蛋白质的核酸, 并首先使用了核酸这一名词。从此以后, 许多化学家对核酸的组成进行了研究, 发现核酸是由有机碱、糖和磷酸组成的; 其水解产物中最常见的碱基是鸟嘌呤、腺嘌呤、尿嘧啶, 胞嘧啶和胸腺嘧啶; 核酸中的糖大多是戊糖, 分核糖和脱氧核糖两类。但对核酸究竟具有怎样的结构, 其在生物体内的功能是什么仍一无所知。直到 20 世纪 20 年代还把核酸从它制备的来源, 错误地分为“动物核酸”和“植物核酸”两大类。所谓动物核酸是指从动物组织制得的核酸, 它的成份中含有“植物核酸”所没有的脱氧核糖和胸腺嘧啶, 而所谓植物核酸是指从酵母和麦胚这一类植物制得的核酸, 它的成份中含有“动物核酸”所没有的核糖和尿嘧啶。后来进一步的研究才发现无论在动物或植物的细胞核内部都可找到所谓的动物核酸, 而在动植物的细胞质内都可找到所谓的植物核酸。这样就不再有动植物核酸之分了, 而根据核酸中的核糖不同, 分成核糖核酸和脱氧核糖核酸。这样的分类一直沿用至今。1944 年 O. T. Avery 等在研究肺炎球菌的转化现象时, 发现从一种类型的肺炎球菌制备得到的 DNA, 可以引到另一类型的肺炎球菌体内, 并使后者的遗传性状改变成前者。由此得出了遗传物质是 DNA 的结论; 并阐明了 DNA 的重要生物功能, 从而改变了当时认为蛋白质才是遗传物质的错误看法。从此核酸研究的重要性受到了广泛的注意, 不仅促进了生物学家对核酸进行更深入的研究, 而且也引起其它学科科学家的兴趣; 特别是 DNA 究竟具有怎样的结构引起了结构化学家和物理学家的浓厚兴趣(如图1)。

在 20 世纪 50 年代前后, 用 X 射线衍射方法研究

生物大分子的结构工作已经开始。当时对一些 DNA 和蛋白质的纤维衍射图进行了较仔细的分析, Wilkins 等从 DNA 的纤维衍射图推测 DNA 纤维螺旋轴的周期是 34 Å, 螺旋的半径是 10 Å, 每圈螺旋含十个核

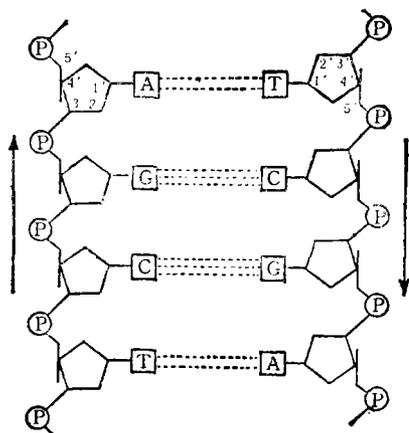


图1 脱氧核糖核酸单链的化学组成

苷酸残基, 磷酸根要处于螺旋的外缘; 从纤维密度考虑, 每一圈螺旋应该大大多于十个核苷酸残基, 因此 DNA 螺旋很可能是三个或两个 DNA 分子构成同轴的螺旋。虽然纤维衍射图仅提供很有限的结构信息, 还不能建立 DNA 的完整的结构模型, 但 Wilkins 已经对 DNA 的结构特点作了基本正确的描述。50 年代初期要想通过单晶体 X 射线衍射法来了解生物大分子的完整的三维结构还是十分困难的事, 因为还没有解决衍射相角问题的方法。但那时结构化学家根据从有关小分子的单晶结构分析所获得的结构信息, 结合结构化学基本原则, 预测了蛋白质分子的结构模型。这就是 Pauling 在 1951 年提出的著名的  $\alpha$ -螺旋结构。Pauling 根据当时已经测定的各种氨基酸的单晶结构, 以及一些小肽的晶体结构资料, 总结出氨基酸内键长键角的数据, 并指出肽键具有平面性的特点, 按照在蛋白质结构中满足最大限度成氢键的原则, 他提出了在蛋白质分子结构中存在由氢键维系的  $\alpha$ -螺旋结构, 并建立了  $\alpha$ -螺旋的结构模型。他的预测不久就被 Kendrew 等人测定的第一个蛋白质分子——鲸肌红蛋

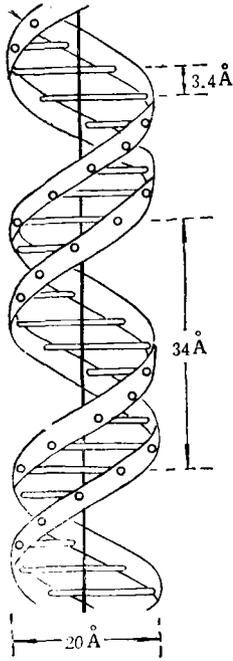


图2 Watson 和 Crick DNA 双螺旋结构模型。两条带表示磷酸根糖环主链,水平棍表示连接两条链的碱基对

他们为此曾获得了 1962 年的诺贝尔生理学医学奖。Watson 和 Crick 提出的双螺旋结构中两条螺旋围绕着同一个轴以右旋的形式盘绕在一起,不考虑碱基间的差别的话,此两条链彼此间有一个垂直于螺旋轴的二重轴相关联,碱基指向螺旋的内部,而磷酸根基团在外部,磷原子离螺旋轴的距离是  $10 \text{ \AA}$ , 糖环粗略地垂直于碱基平面,在每条链的伸展方向间隔  $3.4 \text{ \AA}$  有一核苷酸残基,在同一链上相邻的核苷酸残基之间

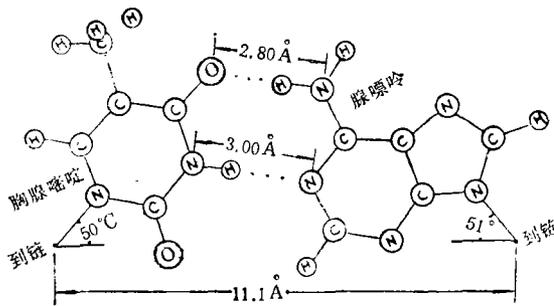


图3 腺嘌呤与胸腺嘧啶配对形成碱基对的大小

的角度是  $36^\circ$ , 所以在一条链中每隔 10 个核苷酸残基, 结构就出现一次重复。这与从 DNA 纤维衍射图上分析所得的结果完全吻合。这一双螺旋结构是通过两条链间碱基对间的氢键联系在一起的, 整个结构好

白的空间结构所证实。人们能否用当时已有的关于 DNA 的信息去预测 DNA 的空间结构呢? 回答当然是肯定的。Pauling 在成功预测了蛋白质  $\alpha$ -螺旋结构后, 也预测了 DNA 的结构。他提出了三条链共轴的螺旋结构, 除碱基对不同外三条链具有三重轴的关系, 磷酸根基团处在螺旋的核心, 碱基转向螺旋的外面。他还给出了这一模型的各原子坐标。但是这一次他没成功。因为带负电的磷酸根基团靠近螺旋轴将引起相互的排斥, 而且在 Pauling 的模型中有些原子间的 van der Waals 距离也太近了。以后还有一些人提出过 DNA 的结构模型, 其中以 Watson 和 Crick 两人提出的 DNA 分子的双螺旋结构为最著名, 如图 2。

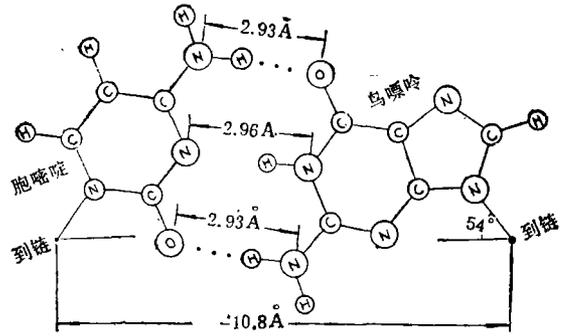


图4 胞嘧啶与鸟嘌呤配对形成碱基对的大小

像是一螺旋状的梯子, 碱基对平面是梯子的踏板。而且这一碱基对是非常专一的, 也就是说一条链上的碱基是腺嘌呤, 则另一条链上碱基必须是胸腺嘧啶与之配对, 反之亦然。类似地鸟嘌呤必须和胞嘧啶配对。只有这种配对方式才能使两种碱基对有差不多的大小。见图 3、图 4。这种专一性的配对是 DNA 双螺旋结构的关键, 它正好解释了 Chargaff 早已发现的 DNA 中碱基之间的一些规律性, 即嘌呤的总数等于嘧啶的总数, 含氮碱基(腺嘌呤和胞嘧啶)的总数等于含酮碱基(鸟嘌呤和胸腺嘧啶)的总数, 腺嘌呤和胸腺嘧啶的摩尔数相同, 鸟嘌呤和胞嘧啶的摩尔数相同。这种专一性的配对还意味着两条链有互补性, 每一条链上的碱基顺序决定着另一条链上的碱基顺序, 说明每一条链均可作为模板以指导另一互补链的合成, 从而产生与亲本相同的双链 DNA 分子, 正如 Watson 和 Crick 指出的那样, 这直接暗示着一种可能的遗传物质的复制机理。在 DNA 双螺旋的模型上可以明显地看到双螺旋的表面形成两个凹槽, 其中一个槽大些也略深一些称大沟, 另一个槽小些浅些称小沟, DNA 和蛋白质的相互识别与沟内的碱基有关。

Watson 和 Crick 的 DNA 双螺旋结构模型把结构研究与功能研究联系起来, 为现代分子生物学奠定了基础, 大大推动了分子生物学的发展; 也促使使用衍射方法测定 DNA 结构的工作在深度和精度上有了进一步的发展。

六十年代用单晶 X 射线衍射方法测定生物大分子的结构日趋成熟, 已经成功地测定了一些蛋白质分子的空间结构。但由于培养 DNA 大分子的单晶还很困难, 因此, 关于 DNA 结构的测定工作还不得限于测定寡聚核苷酸和有关的小分子单晶结构。由此来积累核苷酸中的键长键角数据, 了解碱基对之间成氢键的形式, 以及碱基和糖环的立体化学信息等。根据这些信息结合 DNA 纤维衍射数据, 建立更精确的 DNA 结构模型, 然而根据模型计算衍射强度与实际测得的纤维衍射强度比较。自从 1966 年引入用最小二乘方法来修正原子坐标后, 衍射强度的拟合工作有了极大的

改进,使模型的精度和正确度都有很大提高。当 DNA 钠盐在相对湿度 90% 时, DNA 纤维具有 B-DNA 结构,相对湿度降到 75% 时,就变成了 A-DNA 结构。当相对湿度降到 66% 时且在锂盐的环境下,则具有 C-DNA 结构。到了 70 年代,随着人工合成寡核苷酸的技术日趋成熟,合成指定顺序的小片段 DNA 已成为可能,并且也培养出了小片段 DNA 单晶,因此有一系列的小片段 DNA 的单晶结构分析证实了各种 DNA 结构形式的存在。第一个原子分辨率的 DNA 结构是 1978 年测定的四核苷酸片段  $d(ATAT)$ , 以后相继测定了四核苷酸  $d(CCGG)$  片段,它具有 A-DNA 结

表 1 三种 DNA 模型的结构参数

模型类型	碱基倾角	碱基夹角	碱基升高	螺距	每圈碱基数	小沟宽	大沟宽	小沟深	大沟深
B-DNA	0°	36°	3.37 Å	34 Å	10	5.7 Å	11.7 Å	7.5 Å	8.5 Å
C-DNA	6°	38°	3.31 Å	31 Å	9.3	4.8 Å	10.5 Å	7.9 Å	7.5 Å
A-DNA	20°	32.7°	2.56	28 Å	11	11.0 Å	2.7 Å	2.8 Å	13.5 Å

构;十二核苷酸  $d(CGCGAATTCGCG)$  片段具有 B-DNA 结构。A、B、C 三种 DNA 结构模型的有关参数见表 1。它们的主要区别在于 A-DNA 每圈螺旋含 11 个核苷酸残基,螺距是 28 Å, B-DNA 含 10 个核苷酸残基,螺距是 34 Å; A-DNA 的小沟宽度大大超过大沟宽度,而小沟的深度远不及大沟的深度, B-DNA 中大沟的宽度比小沟大,深度两者相差不大。这样 A 型螺旋比 B 型螺旋拧得更紧一些,碱基倾角大一些。C-DNA 基本上与 B-DNA 相似。自然界的现象是丰富多采的,除了上述三种 DNA 结构外,在一些均聚或

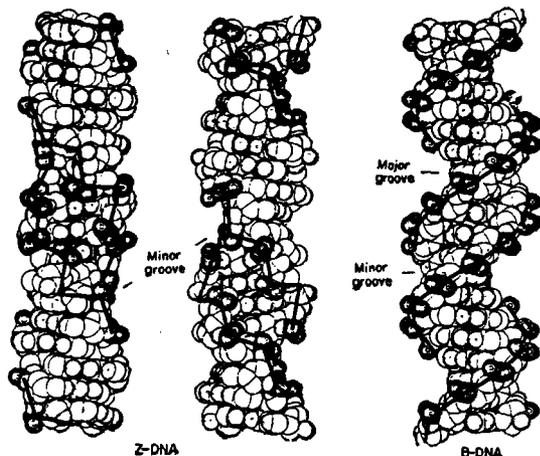


图 5 Z-DNA 和 B-DNA 范氏半径表达的结构模型。左边两个 Z-DNA 模型相互间绕螺旋轴旋转了 30°。黑粗线代表主链骨架。

共聚核苷酸中还发现有与 B 型和 A 型构象略有不同的几种变型,但这些都属于右旋的 DNA 结构。

1979 年 Rich 和他的同事测定了六核苷酸

表 2 B-DNA 与 Z-DNA 结构比较

	B-DNA	Z-DNA
螺旋方向	右旋	左旋
每圈残基数	10	12
螺旋直径	~20 Å	~18 Å
每对残基间轴向升高	3.4 Å	3.7 Å
螺距	34 Å	45 Å
碱基平面倾角	6°	7°
相邻碱基对之间转角	36°	60°
轴心与碱基的关系	穿过碱基对	不穿过碱基对

$d(CGCGCG)$  片段的单晶结构,令人惊奇地发现它具有左手双螺旋结构,因为磷酸根和糖环组成的主链骨架具有锯齿形状,故他们命名为 Z-DNA

(见图 5)。它与右旋的 B-DNA 结构有着十分明显的不同。Z-DNA 每圈螺旋含有 12 个残基,螺距是 45 Å,每两个核苷酸残基作为螺旋中的一个重复单元,而不是 B-DNA 中每个核苷酸残基作为一个重复单元,所以在 Z-DNA 中核苷酸残基重复单元间的转角是 60°, B-DNA 的转角是 36°,在 Z-DNA 中双螺旋表面的大沟已不存在,变成了凸出的曲面(见图 5)。它与 B-DNA 构象不同的关键在于脱氧鸟苷中 ( $dG$ ) 糖苷键是顺式的,脱氧胞苷 ( $dC$ ) 糖苷键是反式的,而在 B-DNA 中两个糖苷键都是反式的。Z-DNA 和 B-DNA 结构差别列于表 2。

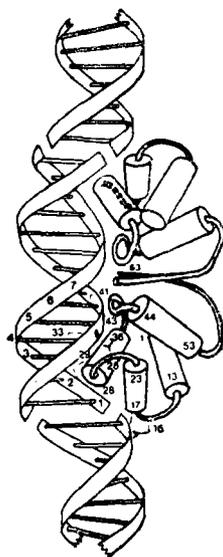


图 6 噬菌体 434 阻遏子和 DNA 复合物的结构示意图。 $\alpha_1$ (残基 28-36)躺在 DNA 的大沟中,并和 DNA 碱基有相互作用。

随着分子生物学的发展,一些 DNA 和 DNA 结合蛋白的复合物单晶结构也已经有所报导。人们想通过这类复合物单晶结构的测定来了解 DNA 和蛋白质的相互作用,以及它们之间专一性识别的机制。但目前已经测定的这类复合物的结构还不多,还不能从中总结出客观的规律。在噬菌体 434Cro 和 DNA 复合物晶体结构及阻遏子和 DNA 复合物晶体结构中,已经发现在结合蛋白中对识别具有重要作用的第三段  $\alpha$  螺旋都躺在 DNA 的大沟中,而且在蛋白分子中的氨基酸残基与 DNA 的碱基之间存在着氢键联系(见图 6),识别蛋白质分子起着重要作用的。