

专论与综述

DNA 的化学合成及其应用

静 国 忠

(中国科学院生物物理研究所, 北京)

通过二十多年的探索和研究, 脱氧核糖寡核苷酸(DNA)片段的化学合成得到了突飞猛进的发展。DNA 固相合成技术的问世更使 DNA 化学合成技术达到了精确、高效和自动化。重组 DNA 技术以及外源 DNA 分子在异源体系中的表达为 DNA 结构功能及基因表达调控机制的研究打开了新局面, 使得生物学的研究发生了革命性的变化, 而 DNA 分子的化学合成成为分子生物学家对特定的基因进行遗传工程的操作, 选择性地对 DNA 分子进行改造提供了崭新的手段。借助于化学合成的 DNA 片段, 人们可以使一个基因在特定的核苷酸位点上产生取代、插入以及缺失等突变, 进而测定这种改变对其生物活性所造成的影响, 这就是当今应运而生的蛋白质工程的基础, 而蛋白质工程则是研究蛋白质结构与功能关系的重要手段, 并为工业生产高稳定性、高效的酶制剂及蛋白质类药物开辟了新的途径。

本文将集中叙述 DNA 片段的固相合成及其在生物、医学研究中的应用, 并根据我们在工作中的经验对有关部分进行较详细的讨论和说明。

一、DNA 的化学合成

DNA 片段的化学合成技术进步的主要标志是高效率的固相合成法取代了通常的液相法。固相法的优点在于偶联效率高、合成产物易于分离纯化、最终产率高且整个操作过程非常简化、易于掌握。目前所采用的 DNA 片段合成的装置有两种: 一种是手动合成仪, 另一

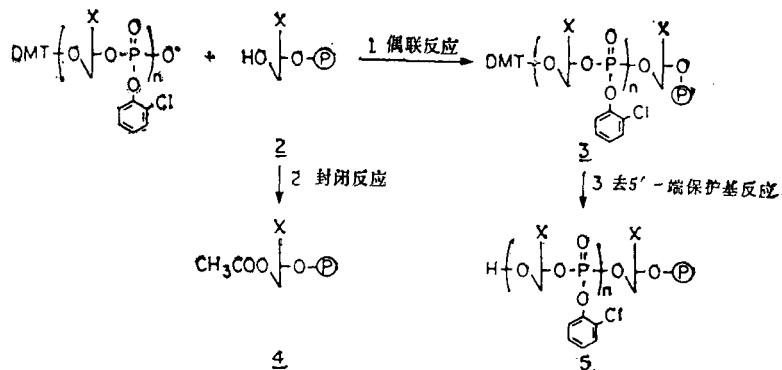
种是自动合成仪。前者经济实惠, 易于掌握及维修, 适于一般实验室应用; 后者昂贵, 然而自动化程度高, 适于大量进行基因合成工作的实验室使用。今后, DNA 合成化学的研究将向着进一步提高偶联效率、缩短偶联时间、在单位时间内合成更长更多的 DNA 片段方向前进。

1. DNA 片段固相合成法的原理及步骤

DNA 片段的固相合成有两种基本的方法, 即磷酸三酯法和亚磷酸三酯法。前者的基本步骤是将活化了的核苷酸, 在存在特定的偶联试剂的条件下依次地同与固相载体相连的核苷 5'-羟基产生偶联反应, 从而在参加反应的两核苷酸之间形成磷酸三酯键。合成的整个过程包括偶联反应、封闭反应、去 5'-端保护基反应。对于亚磷酸三酯法而言, 除了上述几个步骤外还包括一个氧化反应, 使亚磷酸三酯键转化成为磷酸三酯键(图 1)。

DNA 片段化学合成中最难的一步是制备各种带保护基的单核苷酸, 幸运的是这些试剂目前已市售。在磷酸三酯法中, 3'-磷酸三酯化合物(可以是单体、双体或三体)作为基本的偶联单位。偶联反应的关键是整个反应要在绝对无水条件下进行, 各种试剂的纯度将影响合成的 DNA 片段的长度。MSNT(1-(mesitylene-2-Sulfonyl)-3-nitro-1.2.4-triazolide) 作为偶联剂, 其平均偶联效率为 95% 以上。近来, 人们用芳基磺酰氯(arylsulfonyl chloride) 及 N-甲基咪唑(N-methylimidazole) 作为偶联剂, 其偶联效率同 MSNT 相当。聚苯乙烯同 1% 的二乙烯苯的共聚物以及聚二甲基丙烯酰胺的衍生物是常

a) 磷酸三酯法



b) 亚磷酸三酯法

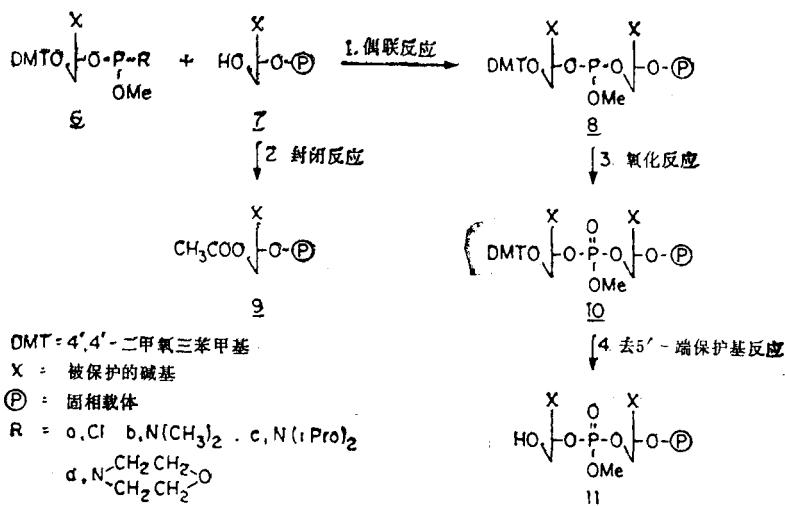


图1 两种DNA片段固相合成法基本过程

a) 磷酸三酯法, b) 亚磷酸三酯法

用的两种固相载体。此外硅胶及有孔玻璃也应用作为固相载体。

每次偶联反应后，在N,N-二甲氨基吡啶的催化下，用乙酸酐将暴露出来但没有进行偶联反应的核苷酸5'-羟基封闭。这样有利于使下一步偶联反应效率维持在高水平。当然，从整个合成过程来说，封闭的主要目的是避免更多不同长度的小分子寡核苷酸片段副产物的生成。

在每次偶联反应之后，寡核苷酸的5'-端总是被DMT-保护基所保护。在磷酸三酯法中，溴化锌的二氯甲烷-丙醇溶液是较理想的去5'-DMT-保护基试剂。然而，如果用聚丙烯酰胺衍生物作为固相载体时，不能用这种溶液作为去DMT-保护基试剂。苯磺酸及三氯乙酸也作

为去DMT-保护基试剂，但常常引起去嘌呤化及异聚化等副反应，特别是合成在3'-端具有脱氧腺苷的寡核苷酸时更要小心。

亚磷酸三酯法是另一种DNA片段化学合成法。根据同亚磷酸基团相结合的反应性基团的不同，又可分为氯亚磷酸三酯法和二异丙基酰胺亚磷酸三酯法。由于氯亚磷酸衍生物本身的不稳定性，使得氯亚磷酸三酯法的应用受到限制。二异丙基亚磷酸酰胺法是亚磷酸三酯法中更现代的方法。同氯亚磷酸衍生物不同，同亚磷酸基团结合的不是氯原子而是一种次级胺（最常用的是二异丙基胺）从而形成二异丙基亚磷酸酰胺化合物。此功能基团的本身并不同羟基产生反应，然而在弱酸性的条件下（如四唑存在下）可转化成四唑的衍生物，这个中间产物能

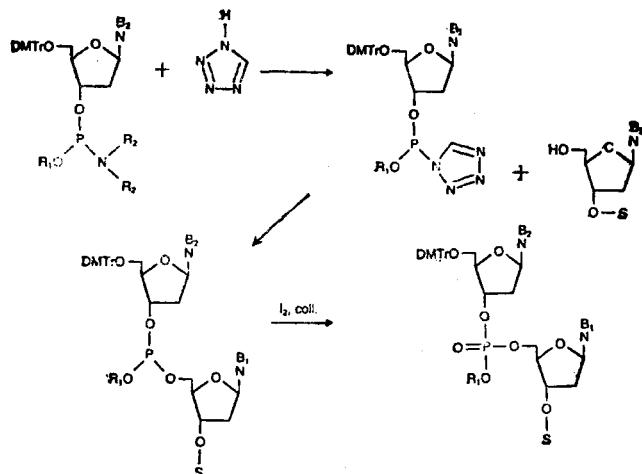


图 2 二异丙基酰胺亚磷三酯法基本过程

很快同另一核苷上的 5'-羟基反应，其偶联效率可达 98% 以上(图 2)。二异丙基亚磷酸酰胺法的最大优点是这种酰胺化合物在粉末状态和溶液状态都十分稳定，有的甚至可用硅胶层析进行纯化。另一大优点是偶联反应速度很快，向寡核苷酸上加一个核苷酸只需十分钟。亚磷酸三酯法目前已广泛用于 DNA 自动合成系统。

2. 化学合成 DNA 片段的分离和纯化

从固相合成中得到的最终产物是一个混合物，即除了全长 DNA 片段外还包括一系列不

同长度的寡核苷酸片段。DNA 片段合成后第一步是将合成的寡核苷酸片段从固相载体上切下并去除各种保护基。第二步是通过各种方法得到纯净的全长 DNA 片段。

(a) DNA 寡核苷酸片段上各种保护基的去除：为了避免副反应以及增加偶联单位的溶解度，合成时几乎所有的核苷酸的功能基团都曾被保护起来，在合成后必须要把这些保护基团去除。两种 DNA 片段固相合成法在这一步是非常相似的(图 3)。

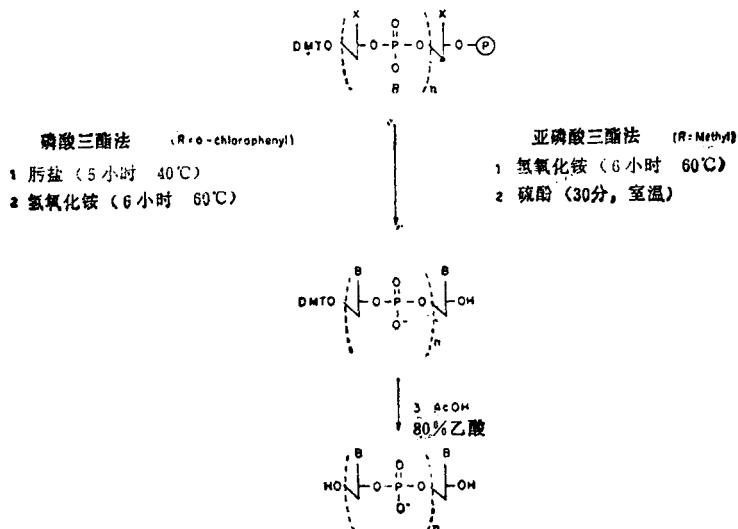


图 3 去除 DNA 寡核苷酸上全部保护基的过程

X: 被保护的碱基
R: 磷氯苯基或甲基
B: 碱基 (A, C, G, T)

· 磷酸基团：在磷酸三酯法中保护磷酸基团的磷氯苯基用 1,1,3,3-四甲基胍-2-硝基苯甲醛去除。正常的磷酸三酯键并不为这种试剂所破坏。肟盐的另一种作用是使在偶联反应中被修饰的鸟便嘌呤再生成鸟便嘌呤。在二异丙基亚磷酸酰胺法中起保护作用的甲基，通常用硫酸处理后去除。

氨基：碱基杂环上的保护基主要是保护碱基上的氨基，如在腺嘌呤和胞嘧啶上的苯甲酰基、鸟便嘌呤上的丁酰基，可用 30% 的浓氢氧化铵溶液在 60°C 处理 6 小时去除。

5'-羟基：寡核苷酸 5'-羟基上的 DMT-保护基在最后用 80% 的乙酸处理去除。如选择 HPLC 法纯化，DMT-保护基在纯化后去除；如用聚丙烯酰胺凝胶电泳法纯化，在纯化前或纯化后去 DMT-保护基都可。为将去嘌呤化作用减到最低限度，处理时间最好不要超过二十分钟。

(b) DNA 寡核苷酸片段的纯化

有两种基本的方法用于 DNA 寡核苷酸片段的纯化，一种是层析法，一种是凝胶电泳法。

层析法又分为 (i) 高压液相层析：如果需要量在 40 微克以下时，采用此法可得到高纯度的样品。然而，用高压液相层析只能对 25 到 30 个核苷酸长度的 DNA 片段进行纯化，对较长的 DNA 片段而言，用此法是非常费时的。(ii) 阴离子交换柱层析：离子交换树脂已广泛地用于像多核苷酸类的荷电化合物的纯化。在离子交换柱上，全长的产物总是比短的核苷酸片段迟些时间从柱上洗脱下来。人们曾用在 Whatman 公司所生产的 partisil 10 SAX 柱上用含 60% 甲酰胺的磷酸缓冲液进行梯度洗脱，分离寡核苷酸的混合物，可将长度约为 25 个碱基的核苷酸纯化。(iii) 逆相柱层析：理论上，用固相法所合成的寡核苷酸中只有全长分子的 5'-端具有完整的 DMT-保护基团。当利用逆相 μ Bondapak C₁₈ 柱进行层析后，可将它们同各种不具 DMT-保护基团的副产物分离。最后，在除去 DMT-基团后再重复进行逆相柱层析可使产物得到进一步纯化^[1]。

聚丙烯酰胺凝胶电泳法：Maxam 和 Gilbert^[2]第一次发现用变性聚丙烯酰胺凝胶电泳可将只有一个核苷酸差异的 DNA 分子进行分离鉴定。正是根据这一发现，人们将尿素-聚丙烯酰胺凝胶电泳应用于 DNA 寡核苷酸片段的分离纯化。通常用含 ηM 尿素的 20% 聚丙烯酰胺凝胶进行电泳分离，一般用制备胶每次可纯化一百微克的样品。实验表明，在很多情况下在无尿素的聚丙烯酰胺凝胶上，寡核苷酸片段也可以得到很好的分离^[1]。

以上两种方法选用哪种，要根据对 DNA 样品的要求和实验室的实际情况而定。一般说来，化学实验室偏用前者，而生物实验室偏用后者。我们曾用聚丙烯酰胺凝胶电泳纯化人胰岛素基因，人的 γ -干扰素基因片段，所得到的样品完全适用于组建基因的工作。用这种方法纯化的片段，在用 $[\gamma-^{32}P]-ATP$ 经多核苷酸激酶标记后，其比放射活性相当高，可达 $10^9 \text{ cpm}/\text{微克}$ ，是一种很好的标记探针。

3. DNA 寡核苷酸片段的检定

经分离纯化后的 DNA 寡核苷酸片段，应对其质量（包括长度和核苷酸顺序）进行检定。检定 DNA 片段的顺序的方法：(a) Wandering Spot 方法：适用于寡核苷酸长度在 20 个碱基以下的片段^[3]。寡核苷酸片段 5'-端经 $[\gamma-^{32}P]-ATP$ 用 T₄ 多核苷酸激酶标记后，经蛇毒磷酸二酯酶进行部分水解，标记混合物于 pH3.5 的条件下在醋酸纤维纸上进行电泳，然后在 DEAE-纤维素板上进行同系层析，寡核苷酸的顺序通过这一标记消化产物的特定迁移率来测定。此方法的好处在于不仅可测出寡核苷酸的顺序，也可以检出不纯物的污染。(b) 利用 Maxam 及 Gilbert 化学裂解法对 DNA 核苷酸顺序进行分析。纯化后的 DNA 片段 5'-端经放射标记后，将其进行特定的化学裂解，部分裂解的产物按其大小在 DNA 序列胶上进行分离，其核苷酸碱基顺序从凝胶的放射自显影图谱上读出。同(a)法相比，此法对于合成产物中少量的不纯物（~10%）是检测不出来的。(c) 测定合成产物长度的方法：以上两种方法都是

直接测定 DNA 的碱基顺序，我们在工作中建立起一个间接测定顺序的方法。要点是用 DNA 序列胶对纯化后的 DNA 片段的长度进行测定，根据我们的经验，只要长度正确，其碱基顺序是可靠的。此方法的关键是制备出好的分子量标记化合物，我们用 oligo-dT₁₂₋₁₈ 作为起始物质，在末端脱氧核苷酰转移酶的催化下得到 12 到 100 个碱基大小的分子量标记化合物^[4]。

化学合成的 DNA 片段经上述分离、纯化及质量鉴定后，即可用于基因重组、筛选探针、DNA 合成的引物，临床诊断等各种目的。

二、化学合成 DNA 片段的应用

1. 基因的合成

DNA 片段合成技术的发展使得人们可在相当短的时间内合成足够的 DNA 片段，用以组建人工合成的基因，特别是那些用其它技术不易分离的基因。目前已有不少的有用基因通过化学合成而获得，如人的生长激素抑制素、人的生长激素、人的干扰素、人的胰岛素以及人的尿抑胃激素等。基因合成的方法一般而论可大致分为以下两类。

(a) 基因片段的全化学合成：所谓基因片段的全化学合成是首先合成组成一个基因的所有片段，相邻的片段间有 4—6 个碱基(按具体情况而定)的重叠互补，在适当的条件下(主要是温度条件)经退火后，用 T₄ DNA 连接酶将各片段以磷酸二酯键的共价键形式连接成一个完整的基因。因为化学合成的 DNA 片段在纯化后其 5'-端及 3'-端都为羟基，所以在组建基因之前要将 DNA 片段的 5'-端磷酸化。一般地说，处于基因两端的寡核苷酸片段的 5'-端不进行磷酸化，以防止基因本身在 DNA 重组时自身环化。对于较大的基因，一般将基因分成几个亚单位进行分子克隆，然后分离纯化这些亚单位，再重新组成一个完整的基因，最后克隆进适当的表达载体进行表达(图 4)。

(b) 基因的化学一酶促合成：此法的特点是不需要合成组成完整基因的所有寡核苷酸片

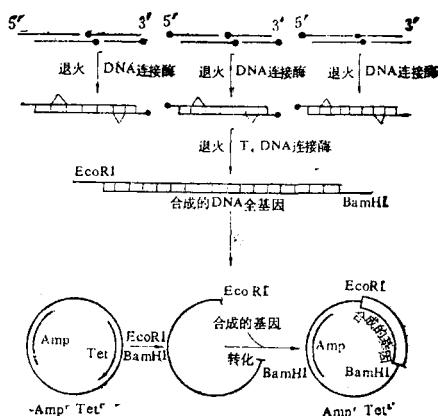


图 4 基因的合成及扩增的过程

段，而是合成其中一些片段，相邻的 3'-末端有一短的顺序互补，在适当的条件下通过退火形成模板——引物复合分子 (template-primer complex)，然后在存在四种 dNTP 的条件下，用大肠杆菌 DNA 多聚酶 I 或反转录酶去填补互补片段之间的缺口，从而形成一个完整的基因(图 5)。同(a)法相比，此法的优越性在于不必合成构成基因的所有片段，特别是由于固相合成法的改进，人们可合成较长的寡核苷酸片段，这样用酶法去填补缺口就可以避免由于合成所有寡核苷酸片段所需过多的时间和费用。用此法我们曾将人胰岛素 A、B 链基因的六个寡核苷酸片段退火后，用 *E. coli* DNA 多聚酶 I 的大片段 (Klenow fragment) 在存在四种

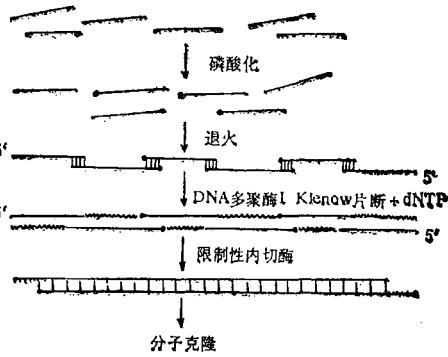


图 5 化学-酶促合成基因的图解

dNTP 的条件下填补片段间的缺口，最后用 DNA 连接酶 (*T₄* DNA Ligase) 连接，得到人胰岛素 A、B 链基因。

2. 基因表达的调控研究

为了研究这一问题人们用化学法合成了各种操纵子，来研究基因转录的调控。通过对大量的核苷酸顺序的比较，人们发现在多数原核生物基因转录起始点的上行区有两个特定的核苷酸顺序同基因转录的起始效率有关，这就是-10及-35区，其碱基顺序分别是TATAATG和TTGACA。如果一个操纵子所含的两个顺序同上述是精确一致，即代表RNA多聚酶最完

美的结合顺序，那么由这两个顺序所组成的启动子就是一个强的启动子。据此人们可设计合成最佳的启动子，用以研究它们同RNA多聚酶之间的相互作用。正是利用这一原则将Lac操纵子的启动子-10的顺序同Trp操纵子的-35顺序结合到一起，从而合成一个杂种启动子Tac（图6）。同较强的trp启动子相比 Tac启动子之转录效率提高了3—5倍。

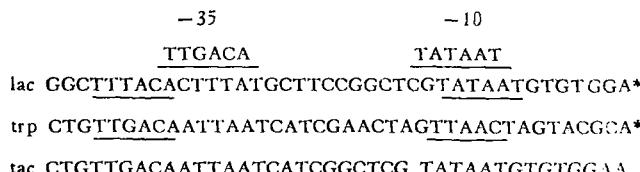


图6 强启动子 Tac 的组建和顺序

E. Jay 等人用化学合成的转录和翻译的调控碱基顺序去研究外源基因在大肠杆菌中的有效表达¹⁵。他们发现，合成的核糖体结合顺序(RBS)同翻译起始密码子 ATG 之间的核苷酸数目影响翻译产物的产率。如图7所示，当 RBS 和 ATG 之间存在 8 个核苷酸时，无蛋白质产生(t75 克隆)；当缺失一个(t23 克隆)和二个(t1 克隆)核苷酸时产生低表达。当缺失 3 个(t24 克隆)或 4 个(t12 克隆)核苷酸时产生高表达。上述实验指出，在 RBS 和 ATG 之间如果存在一个多余的核苷酸可使翻译效率急剧地降低。

(a) 改进 M_{13mp} 系列，用合成的具有多种限制内切酶识别位点的 DNA 顺序 (polylinker) 插入到 M₁₃ 载体上，形成更好的克隆载体，以便将含各种限制内切酶识别位点的 DNA 片段重组于载体上进行顺序分析。例如 M_{13mp2} 只有一个 EcoRI 识别位点，而 M_{13mp18,19} 则含有 EcoRI、SacI、KpnI、XmaI、SmaI、BamHI、XbaI、SalI、AccI、HincII、PstI、SphI、HindIII 等 13 个限制内切酶识别位点。(b) 合成顺序引物 (sequencing primer)，一般引物的长度为 15 或 17 个碱基。虽然目前有市售的通用引物 (universal primer)，然而，有时需要自己合成特定顺序的引物。

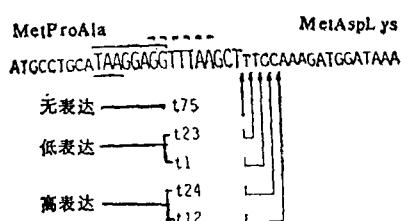


图7 核糖体结合位点(RBS)同 ATG 起始密码子之间碱基数目对表达的影响

3. DNA 的顺序分析

DNA 寡核苷酸片段的合成也应用于DNA顺序的分析中，这是由于 Sanger 发展了双脱氧法测定 DNA 顺序之后，用 M_{13mp} 系列的载体进行 DNA 顺序分析的工作。寡核苷酸片段用于 DNA 顺序分析主要有如下两方面的工作：

4. 基因的分离

由于化学合成 DNA 寡核苷酸片段技术的进步以及分子杂交技术的日益完善，用合成的寡核苷酸作为探针进行基因分离和鉴定已发展成为一项成熟的技术。(a) 分离纯化 mRNA 的亲合试剂。化学合成的 oligodT 结合到纤维素上制成 oligodT 纤维素，利用亲合层析的手段分离 mRNA 来制备 cDNA 基因库，这可能是寡核苷酸最早的应用之一。(b) 用寡核苷酸作为杂交探针。在适当的条件下，当同含有互补顺序的 DNA 或 RNA 分子杂交时，寡核苷酸将与其形成稳定的双链结构，利用这一特性从基因库中分离特定的基因。寡核苷酸探针的设计依赖于已知的或从蛋白质顺序中所推演出的互补分子

的核苷酸顺序。这些化学合成的寡核苷酸探针可用于菌落的原位杂交或基因库的噬斑杂交。特别是 William, I. W. 等人^[6]报道了一个用寡核苷酸片段为探针对高度复杂的基因库进行筛选的方法后，大大地提高了寡核苷酸探针的灵敏度，使此技术得到更普遍的应用。(i) 独特顺序的寡核苷酸探针：即通过各种方法查明某基因的一段核苷酸顺序，将它用化学合成，作为探针去分离、鉴定特定的基因。如 Stewart 和 Sherman 等人^[7]在分析了几种移码突变 (frame shift mutants) 的氨基酸顺序的基础上，推导出酵母细胞色素 C 基因中部分的核苷酸顺序，并以此为探针筛选出含有细胞色素 C 基因的克隆 DNA。然而从一个特定的蛋白质氨基酸顺序推导出一个独特的核苷酸顺序的几率是很低的。近来 Jaye 等人^[8]合成一个非常长的寡核苷酸探针去克隆人的第九因子 (human factor IX) cDNA。合成这个探针的策略是选择出一段由 18 个氨基酸组成的肽段，根据其氨基酸顺序选择出为这些氨基酸编码的最佳密码。如果可能出现错配的话，选择 G-T 而不选择 A-C；用这样的策略合成一个同所要筛选的基因有 80% 以上同源性的探针，在适当的杂交条件下通过杂交筛选出人的第九因子基因。这种单一的独特顺序的探针也已经用于鉴定含有已知 DNA 顺序的克隆，Schulze 等人^[9]用这样的探针从小鼠的主要组织相容性复合体 (major histocompatibility complex of the mouse) 的多基因家族 (由 20—40 个基因组成) 分离出 H-2K^b 基因。(ii) 混合顺序的寡核苷酸探针：基于单碱基对错配对寡核苷酸杂交行为影响的研究，Wallace 等人^[10]提出，合成一个寡核苷酸混合物，它代表了一个特定肽段中氨基酸残基的所有可能的编码顺序，然后用这样的一个寡核苷酸混合物作为探针去

检定被克隆的 DNA 应当是可行的。实验指出，当用 11、14、17 个碱基长的寡核苷酸片段同 $\phi \times 174$ DNA 进行杂交时，即使只有一个碱基对错配，它们所形成的杂交双链结构就显得很不稳定。因此，可以利用这种热稳定性的差异，通过选择适当的杂交和冲洗温度，排除由错配探针同杂交 DNA 片段之间所形成的双链，而仅保留完全匹配的杂交双链复合体，从而筛选出特定基因克隆。这一方法的优越性在于只要我们知道某一蛋白质的一个区段的氨基酸顺序，我们就可以合成一个混合探针从基因库中筛选出特定的基因。目前已有 25 个以上的不同 cDNA 克隆就是用这一方法鉴定出来的。

如何去合成一个混合的寡核苷酸探针，这是化学合成中的一个策略问题。因为我们不可能一个个地合成一个肽段氨基酸的所有组合。为着合成一个混合的探针要有如下两步：①根据密码的简并原理，找出一段肽链氨基酸残基的所有编码可能性(图 8)。②用固相法合成寡核苷酸片段混合物。在合成过程中于这顺序的适当位置上加上偶联单位混合物，使得混合物中的每个顺序的含量大约相等(图 9)。

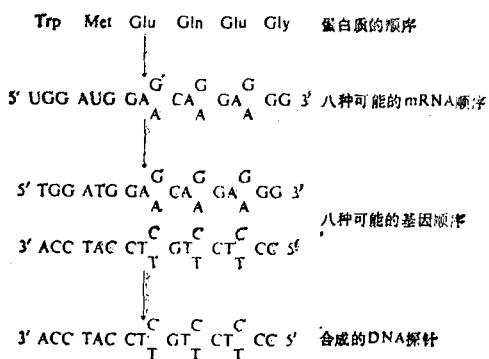


图 8 根据密码简并性的原理推导出的各种可能性的编码

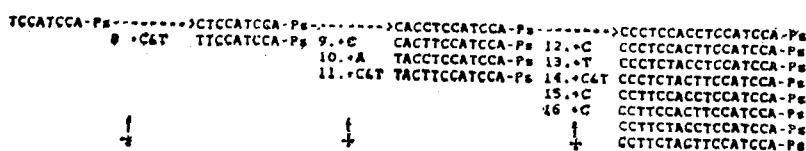


图 9 固相法合成寡核苷酸混合物探针的程序

P_s 为固相载体

(c) 寡核苷酸作为引物用以分离特定的基因:

一般有两种方法: (i) 用寡核苷酸作为引物, 以 mRNA 为模板去制备放射性标记的互补 DNA 链 (primer extension product), 然后以此放射性标记的顺序为探针通过分子杂交去分离检定克隆的 DNA; (ii) 将单链互补 DNA 本身再加倍成双链 cDNA, 然后将此基因拷贝重组入克隆载体进行克隆。无论哪种情况, 同引物互补的 mRNA 的顺序都必须知道或可以从所提供的氨基酸顺序推导出来。作为用逆转录酶以 mRNA 为模板合成放射性标记的杂交探针或特定的基因的引物, 同上述的杂交探针一样也可以有专一的引物 (unique primer) 或含有各种可能性顺序的混合引物 (oligonucleotide primers of mixed sequences)^[1]。

用合成的寡核苷酸作为杂交探针从各种基因库(如 genome 或 cDNA Library) 或细菌克隆中通过分子杂交分离基因和用寡核苷酸引物通过逆转录酶在 mRNA 模板上合成基因这是用以分离基因的两种通用方法。相比之下前者有其优越之处: 第一, 用杂交探针方法有更大的特异性, 这是因为在分子杂交实验中通过选择不同的条件, 我们可以排除即使一个碱基对错配的可能性, 而用引物法至少从技术角度讲是很难做到的。第二, 用引物法从 mRNA 模板上合成的探针受 mRNA 量的影响, 对于在细胞内含量很少的 mRNA, 用这种方法很难得到足够的探针, 而用化学法合成寡核苷酸探针就不存在这样的困难。Smith^[11] 已对用寡核苷酸作为探针进行分子杂交的原理做了论述, 其中包括探针的长度、杂交溶液的离子强度、探针顺序互补性的程度以及杂交温度对杂交双链的稳定性的影响等等, 这些研究为更好地利用分子杂交分离基因以及本文后面将提到的应用分子杂交探针对某些疾病进行诊断等奠定了基础。

5. 基因突变及蛋白质结构功能的研究

实际上基因突变及蛋白质结构功能的研究是一个问题的两个方面, 这是由于蛋白质结构基因的核苷酸顺序的改变将导致于其所编码的

蛋白质分子结构和功能的改变, 蛋白质工程正是利用这一基本原理, 通过改变蛋白质分子编码的结构基因中核苷酸的组成、排列方式, 从而得到一些同野生型蛋白质分子不同的, 具有某些新特性的突变蛋白质分子, 最后通过对野生型和突变型蛋白质的各种特性进行比较, 研究蛋白质的结构与功能的关系。

如何用合成的 DNA 寡核苷酸片段来使基因产生突变? 这里将谈两个基本的方法:

(a) 寡核苷酸指导下的特异性位点突变: 现在, 这一工作多在 M₁₃ 噬菌体上进行, 其要点是合成一段在特定位点上碱基顺序已被改变了的某基因的寡核苷酸片段, 用此片段同含有正常基因的野生型 M₁₃ 单链 DNA 杂交, 并以其作引物在大肠杆菌 DNA 多聚酶 I (Klenow 片段) 的催化下合成 RF DNA, 这样一条链是野生型, 另一条链是突变型, 经分离和增殖后将突变了的基因从 M₁₃ 载体上切割下来重组入表达载体, 从而产生突变蛋白质用于蛋白质结构和功能的研究。正是利用这一方法将干扰素 β 的第一位的半胱氨酸的密码子变成丝氨酸密码子, 使大肠杆菌中 β -干扰素的效价提高了十倍(图 10)^[12]。

为了提高特异性点突变的效率, 人们通常应用较长的突变引物 (mutagenic primer), 这是因为引物长, 其同靶子顺序所形成的杂交双链更稳定且减低了在载体上其他区段产生错误引导 (mispriming) 的可能性。另一种提高突变效率的方法是双引物法, 即同时用两个引物, 一个是带有位点突变的寡核苷酸片段, 一个是同模板完全互补的片段, 第一个引物的位置紧接第二个引物之后, 在适当的条件下进行如上所述的同样操作过程获得具有位点突变的基因。

用 M₁₃ DNA 载体作为特异性位点突变的局限之处在于插入超过 1,000 个碱基对的 DNA 片段就不稳定, 在病毒增殖时易产生缺失。此外, 为了对突变基因的生物活性进行监测, 必须将此基因再克隆进适当的表达载体, 所以人们提出用质粒 DNA 载体作为突变的模板^[13]。

(b) 基因片段取代法: 这一方法的要点是利用在蛋白质结构基因顺序中所存在的限制性

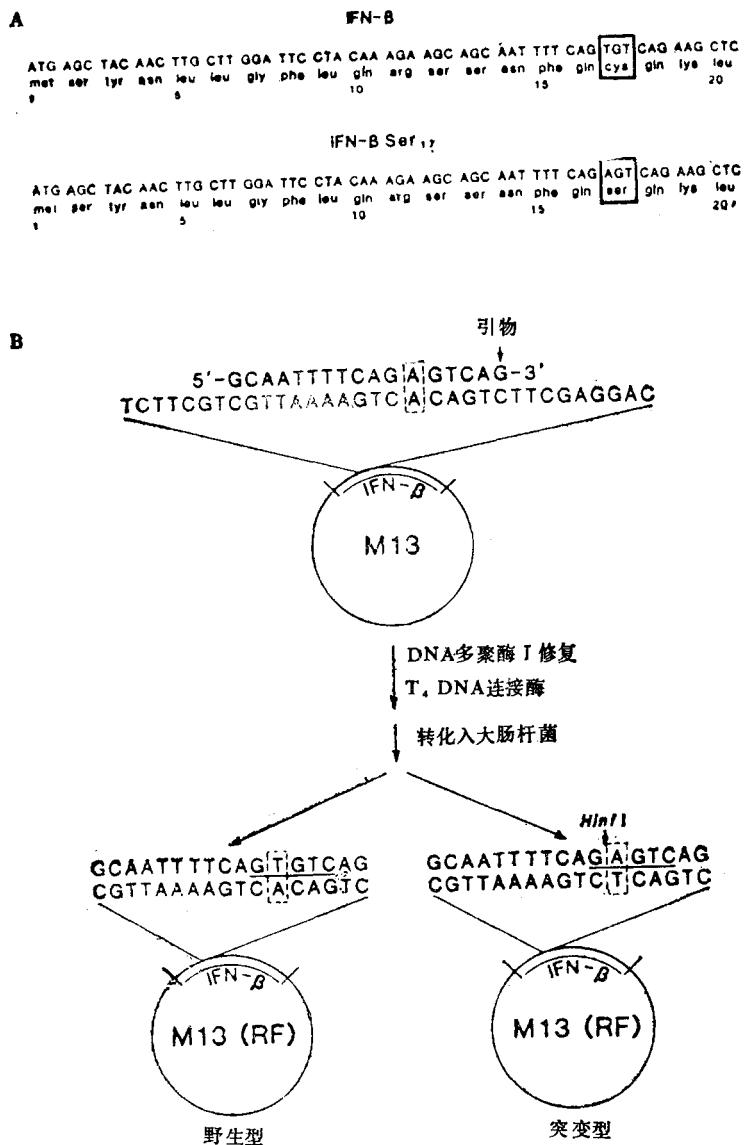


图 10 M₁₃ 噬菌体 DNA 产生特异性点突变示意图
 β -干扰素中的一个碱基由 T 变为 A, 使半胱氨酸变为丝氨酸

内切酶的位点, 用适当的酶去切割 DNA 分子, 从而去掉我们想要改变的那一段基因顺序, 然后插入一段所需的突变顺序(或使其缺失), 从而产生突变的 DNA 分子(图 11)。用此法可同时产生多位点突变, 且由于用凝胶电泳预先去除不需要的野生型基因片段, 所以重组后的野生型本底大大减少, 便于对突变基因进行筛选。然而, 对一个天然蛋白质基因而言, 所含可供利用的限制酶位点相当少, 不足以使我们随意地进行基因片段的取代操作; 为了解决这一问题

人们利用了遗传密码的简并性, 在不改变蛋白质氨基酸顺序的前提下改变某些核苷酸的顺

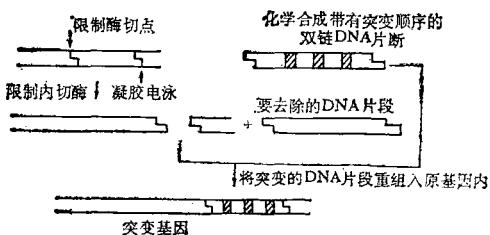


图 11 用基因片段取代方法产生基因突变体示意图

序，从而得到足够满意的限制酶切位点。然而值得注意的是，给定基因的天然核苷酸顺序的改变有时影响这个基因的表达效率。

6. 限制性位点的工程——人工接头的合成

在基因克隆的过程中，有时在一个给定基因的两端或克隆载体 DNA 上没有合适的限制内切酶的位点。利用化学合成去合成自我互补的单链寡核苷酸片段，在一定的条件下通过自我退火形成含有一个限制内切酶位点 (Linker) 或多个限制酶位点的人工接头 (adaptor)，将这些接头通过 T₄ DNA 连接酶接到给定基因或克隆载体 DNA 上，经相应的限制内切酶切割后可对给定的基因进行分子克隆。

7. 疾病诊断探针

利用寡核苷酸片段作为特异性探针对遗传病进行检测已取得可喜的进展。限制内切酶片段从琼脂糖胶转移到硝基纤维素的技术^[13]以及直接的胶杂交技术^[14]为利用特异性的寡核苷酸为探针进行遗传病诊断奠定基础。从理论上讲，即使由于单个碱基取代所产生的遗传病都可以用这一技术检验出来。这一技术目前已用于地中海性的 β^0 -贫血症^[15]、 β^+ -贫血症^[16]以及 α_1 -抗胰蛋白酶缺乏症的诊断^[17]。近来，不少实验室开始用人工合成的探针对一些病毒性感染疾病进行诊断，以及植物病毒病的诊断及防治。

综上所述，人工合成 DNA 目前已发展成为现代生物学研究的有用工具，DNA 化学合成它与分子克隆技术相结合将使分子生物学、医学以及农业、发酵工业的面目出现日新月异的

变化，这将为进一步解决生物学及医学中的重大理论问题做出贡献。

参 考 文 献

- [1] Itakura, K. et al.: *Ann. Rev. Biochem.*, **53**: 323, 1984.
- [2] Maxam, A., Gilbert, W.: *Methods Enzymol.*, **65**: 499—560, 1980.
- [3] Rossi, J. J., Ross, W. et al.: *J. Mol. Biol.*, **128**: 21—47, 1979.
- [4] Jing, G. Z. et al.: *Anal. Biochem.*, 1986. in press.
- [5] Jay, E. et al.: *Biotechnology: Applications and Research*, 388—400, 1984.
- [6] William, I. W. et al.: *Proc. Natl. Acad. Sci. U. S. A.*, **82**: 1585—1588, 1985.
- [7] Stewart, J. W., Sherman, F.: In *Molecular and Environmental Aspects of Mutagenesis* (ed. L. Prakesh, et al.), pp 102—7 Springfield, III: Thomas 1974.
- [8] Jaye, M. et al.: *nucleic Acid Res.*, **11**: 2325—2335, 1983.
- [9] Schulze, D. H. et al.: *Mol. Cell. Biol.*, **3**: 750—755, 1983.
- [10] Wallace, R. B., et al.: *Nucleic Acids Res.*, **9**: 879—894, 1981.
- [11] Smith, M.: In *Methods of DNA and RNA Sequencing* (ed. S. M. Weissman), Praeger publisher, New York, N. Y. 26—68, 1983.
- [12] Mark, D. F. et al.: *Proc. natl. Acad. Sci. U. S. A.*, **81**: 5662—5666, 1984.
- [13] Southern, E. M.: *J. Mol. Biol.*, **98**: 503—517, 1975.
- [14] Conner, B. J. et al.: *Proc. natl. Acad. Sci. U. S. A.*, **80**: 278—282, 1983.
- [15] Piratsu, M., Kan, Y. W. et al.: *N. Engl. J. Med.*, **309**: 284—287, 1983.
- [16] Orkin, S. H. et al.: *J. Clin. Invest.*, **71**: 775—779, 1983.
- [17] Kidd, V. J. et al.: *Nature*, **304**: 230—234, 1983.

[本文于 1986 年 2 月 1 日收到]