

叶寅 王苏燕 田波 主编

核酸序列测定

——
实验室指南

科学出版社

Q52-33
YY

V512C12

核酸序列测定——实验室指南

叶寅 王苏燕 田波 主编

科学出版社

1995

(京)新登字 092 号

内 容 简 介

本书是国内第一部系统介绍核酸序列分析的原理和方法的著作。全书共分 12 章,内容包括核酸序列分析的背景知识、DNA 序列测定的策略、DNA 和 RNA 序列测定的原理和各种实用的实验程序、核酸序列的自动化测定以及核酸序列的数据处理等。本书由具有多年工作经验的专家编写,他们结合国内现有的工作条件和工作经验,着重介绍核酸序列分析的各种实验方法,且实验方法程序化。全书内容丰富,选材代表了该技术的新成就和发展方向,理论联系实际,有较强的理论性和实用性。可作为分子生物学、生物化学、分子遗传学、微生物学以及生物技术等诸领域的科研人员和大专院校师生的实用参考书。

核酸序列测定——实验室指南

叶 寅 王苏燕 田 波 主编

责任编辑 范淑琴 吴瑰琦

科学出版社出版

北京东黄城根北街 16 号

邮政编码:100717

北京朝阳大地印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1995 年 8 月第 一 版

开本: 787×1092 1/16

1995 年 8 月第一次印刷

印张: 21

印数: 1-1 500

字数: 479 000

ISBN 7-03-004518-1 / Q · 556

定价: 38.00 元

前 言

1986年美国能源部(Department of Energy, DOE)资助建立了人类基因组研究中心(Center of Human Genome Studies); 1988年美国国家卫生署(National Institute of Health, NIH)成立了人类基因研究办公室,并于1989年扩大为国家人类基因组研究中心(National Center for Human Genome Research),这两个组织均对人类基因组研究给予资助与规划。随后英、法、日等国相继设立了专门机构和拨出经费,联合国和欧共体也建立了类似的机构,其目的都是推进人类基因组研究。那么,人类基因组研究的目标是什么?为何如此兴师动众?人类基因组研究的目标是测定并分析人类基因组DNA全部核苷酸顺序及结构(人类基因组DNA长达 3×10^9 bp!)、识别所有基因的编码、测定基因的位置及其功能。正如这一工作的组织者所预言的,如果这一巨大工作取得重要突破的话,其研究成果将对生物学及医学有着深远影响。然而这一研究的首要任务、而且是最为繁重的工作则是人类基因组DNA的全序列测定。而 3×10^9 bp无异是一个天文数字,难怪如此兴师动众!

这一前所未有的巨大国际协作攻关工程的提出与发展,至少说明了以下三个方面的问题。其一,核酸序列测定这一在核酸酶学和生物化学的基础上建立并发展起来的技术,是生物学研究发展到一定阶段的产物和需要。核酸是生命的遗传载体,遗传信息以核酸的基本组成单位——核苷酸的不同排列顺序编码在核酸分子中。核苷酸排列顺序改变意味着它的生物学含义的改变。因此,测定与分析核酸序列是研究遗传物质结构与功能相互关系的前提,是分子生物学的基本课题之一。随着分子生物学、分子遗传学、生物化学等前沿学科的迅速发展,核酸序列测定技术有着越来越重要的地位。特别是在分子生物学、分子遗传学等学科的基础上发展起来的基因工程则更是离不开这一技术。其二,核酸序列测定技术从提出与建立之日起,经过不断发展与完善到目前已成为一门相当成熟的研究手段,可以适应生物学研究的更高要求。其三,核酸序列测定技术的发展很快,必将提出更为简便与可靠的方法以适应生物学的研究需要。举人类基因组研究计划为例,人类基因组长达 3×10^9 bp,按目前通用的Sanger方法,一次实验中只能读出400—500bp,好的结果可读得1kb,那么至少需获得 3×10^6 个克隆才有可能获得全长序列(难度远不止如此)。可以想象这里面的工作量和繁杂程度了。因此,必须在测序技术上有新的突破。国际上一些大公司已联手研究开发新的测序技术。

基于核酸序列测定这一技术在生物学研究中广泛而重要的应用以及这一技术的迅速发展,同时考虑到国内对这一技术的认识、掌握与应用相对落后,而且国内尚无系统介绍核酸序列测定的专著,因此将这一技术全面而系统地介绍给国内的同行们,使这一技术更好、更广泛地应用于国内生物学的研究是非常必要的。这也是作者们的最大愿望。

核酸分为两大类:DNA和RNA。首先建立与发展起来的是RNA序列测定技术。然而,早期建立的RNA序列测定方法——小片段重叠法——相当费力与费时。例如Holley及其同事为测定酵母丙氨酸tRNA的核苷酸序列(75个核苷酸),花了9年时间,共用去

了140kg酵母,从中提取200g tRNA,最后纯化出1g 丙氨酸 tRNA。显然,以如此的人力、物力以及时间耗费,要进行核酸序列分析并进而进行分子生物学的深入研究,其难度可想而知。而且,此时,DNA(主要的遗传物质)的序列测定技术远不如RNA的序列测定发展得快。幸运的是,在1975年研究者们终于突破了用于RNA或蛋白质序列测定的策略——小片段重叠——这一框框的束缚,发展了一种新的策略——加减法。在这一策略中,不直接测定DNA片段中碱基的组成和顺序,而是以测定片段的分子长度来推测碱基的顺序。其后发展起来的各种DNA序列测定的方法,如1977年Sanger提出的双脱氧链终止法,Gilbert等发展起来的化学降解法,都是以这个原理为基础的。这个原理也应用到RNA序列测定中并因此建立了很多RNA序列测定的新方法。

与国外同类书相比,本书内容更丰富与完整,且重点突出。在编写过程中,参照了国外一些书籍的有关内容。但考虑到本书的系统性、完整性以及核酸序列测定技术的新进展,因此采用了一套完全不同于国外同类书的编排布局。全书在第一章(引言)安排了与核酸序列分析有关的背景材料;序列测定的高压电泳系统在一般同类书中或简或略,但考虑到国内读者对这一方面的内容了解甚少或模糊,因此,给予了相当的篇幅。本书的重点之一——DNA测序策略的制订远比其他同类书详细,因为这是相当重要的内容。DNA测序的双脱氧链终止法及Maxam等的化学法则是本书的又一个重点。此外,与同类书相比,聚合酶链式反应(PCR)测序、银染法测序、非放射性标记的核酸序列测定以及DNA自动序列仪测序是本书首先引入的内容。核酸序列的计算机处理占据相当大的篇幅,也是本书的一大特点。

作为一部实验指南,本书除必要的基础理论知识的叙述外,着重点放在详细而实用的实验程序上。根据国内现有的工作条件,结合作者们多年的工作经验,本书给读者介绍能够开展工作的基本实验条件、实用的实验程序、任何实验方法程序化。同时结合国内外的产品介绍,可使广大读者方便地掌握产品信息(根据作者们的经验,将给读者推荐实用而节俭的产品)。作者最大的愿望是,读者一旦拥有本书,即可在本书的指导下开展工作并逐步提高。

需要指出的是,本书凝聚着集体的智慧。参与本书编写的作者有着丰富的实践经验和深厚的理论功底,他们的参与,给本书增添了色彩。

本书在出版过程中得到了各方面的帮助,特别应该提到的是中国科学院微生物研究所周培瑾所长,他非常关心本书的编写工作;王文婷输入了全部文稿,张霞为本书绘制了插图,在此谨致谢意。我们还要感谢中国科学院科学出版基金专家委员会,是他们慷慨资助使本书得以顺利出版。

叶 寅 王苏燕 田 波

1994年5月

目 录

前言

第一章 引言	叶寅(1)
1.1 核酸的组成与结构	(1)
1.2 核酸的变性与复性	(6)
1.3 核酸酶与核酸的降解	(6)
1.4 核酸的磷酸化与去磷酸化	(10)
1.5 DNA 合成与 DNA 聚合酶	(11)
1.6 用于序列测定的载体系统	(14)
1.7 分子克隆	(16)
第二章 序列胶及高压电泳系统	叶寅(18)
2.1 引言	(18)
2.2 电泳装置系统	(19)
2.3 序列胶的制备	(24)
2.4 高压电泳	(29)
2.5 序列胶电泳后的处理	(31)
2.6 从序列胶片上读取 DNA 序列	(33)
第三章 DNA 序列测定——测序策略	叶寅(36)
3.1 引言	(36)
3.2 随机测序——建立随机重叠克隆文库	(38)
3.3 利用 T4 DNA 聚合酶构建嵌套系列缺失突变体	(43)
3.4 利用核酸外切酶 BAL 31 构建嵌套系列缺失体	(66)
3.5 利用外切核酸酶 III 构建嵌套系列缺失突变体	(71)
3.6 利用 DNase I 构建嵌套系列缺失突变体	(77)
第四章 双脱氧链终止法——M13 单链测序系统	王苏燕 叶寅(82)
4.1 引言	(82)
4.2 双脱氧链终止法测序原理	(82)
4.3 M13 系统——单链 DNA 的序列测定	(87)
4.4 单链模板 DNA 的制备	(94)
4.5 双脱氧测序反应	(98)
4.6 双脱氧测序反应和序列胶中出现的问题	(113)
4.7 多样品操作的程序化	(117)
第五章 双链 DNA 的序列测定	王苏燕 叶寅(121)
5.1 引言	(121)
5.2 双链模板 DNA 的制备	(127)

5.3	质粒 DNA 的变性——单链模板的获得	(136)
5.4	测序反应	(138)
第六章	PCR 用于序列测定	王铁辉(142)
6.1	引言	(142)
6.2	对 PCR 双链产物测序	(146)
6.3	不对称 PCR 产生单链测序	(152)
6.4	直接 PCR 测序	(158)
第七章	银染测序	王苏燕(167)
7.1	引言	(167)
7.2	银染测序	(168)
7.3	测序中出现的问题及解决方法	(177)
第八章	非放射性标记的序列测定	范俊(179)
8.1	引言	(179)
8.2	测序反应	(184)
8.3	电泳及转膜	(188)
8.4	显色反应	(189)
8.5	结果举例及注意事项	(191)
第九章	DNA 自动测序技术	沈天翔(193)
9.1	引言	(193)
9.2	自动测序技术原理	(193)
9.3	DNA 模板的制备	(204)
9.4	DNA 聚合酶延伸反应	(207)
9.5	序列胶的制备	(212)
9.6	样品准备和电泳	(213)
第十章	DNA 序列分析——M & G 法(化学降解法)	王苏燕(215)
10.1	引言	(215)
10.2	化学降解法测序原理	(215)
10.3	DNA 样品的制备	(220)
10.4	DNA 样品的末端标记	(221)
10.5	单末端标记 DNA 片段的分离和提纯	(224)
10.6	合成的寡核苷酸 DNA 样品的 5' 末端标记	(228)
10.7	特异性的碱基化学裂解反应	(229)
10.8	合成的寡核苷酸的 M & G 测序法	(235)
10.9	序列胶高压电泳及序列的阅读	(235)
第十一章	RNA 序列分析	周志宏 田华松(236)
11.1	引言	(236)
11.2	RNA 的提取与纯化	(242)
11.3	双向层析法	(249)
11.4	酶法及化学法	(254)

11.5	反转录酶法.....	(266)
11.6	末端鉴定法.....	(273)
第十二章	核酸序列的计算机分析	徐 军 陈润生(276)
12.1	计算机、软件和核酸数据库	(276)
12.2	组装一个新测得的序列.....	(284)
12.3	如何分析一个新测定的 DNA 序列	(291)
附录		(308)
一、	各种不同用途的 <i>E. coli</i> 菌株	(308)
二、	各种培养基及抗生素	(310)
三、	放射性同位素	(313)
四、	核酸的性质	(314)
五、	一些经常使用的有机化合物的理化性质	(318)
六、	纯化核酸的试剂	(319)
七、	酶类及酶用缓冲液	(319)
八、	电泳	(321)
九、	序列载体	(322)

第一章 引言

叶寅

(中国科学院微生物研究所 北京 100080)

1.1 核酸的组成与结构

1.1.1 核酸的组成

核酸是高分子量的多体化合物。自然界存在两种类型的核酸:核糖核酸(ribonucleic acid, RNA)和脱氧核糖核酸(deoxyribonucleic acid, DNA)。核酸完全水解后产生嘧啶碱基(pyrimidine base)和嘌呤碱基(purine base),一种糖类(sugar)及磷酸。部分水解则可产生核苷(nucleoside)和核苷酸(nucleotide)化合物。

核苷酸是核酸的基本组成单位,核苷酸又由磷酸基团、戊糖(核糖)和碱基通过一定的结合方式而构成。组成核酸的糖类分为2种:脱氧核糖(2-deoxyribose)和核糖(D-ribose);而碱基主要包含5种:腺嘌呤(adenine, A)、鸟嘌呤(guanine, G)、胞嘧啶(cytosine, C)、胸腺嘧啶(thymin, T)、尿嘧啶(uracil, U)。

1.1.1.1 碱基(base)

1. 嘧啶碱基:嘧啶碱基是嘧啶化合物的衍生物(图1.1)。核酸中常见的3种嘧啶碱基中,胞嘧啶(C)存在于DNA和RNA两类核酸中,胸腺嘧啶(T)只存在于DNA中,尿嘧啶(U)只存在于RNA中。除此之外,还含有一些稀有碱基,例如在某些大肠杆菌噬菌体核酸中,胞嘧啶由5-甲基胞嘧啶或5-羟甲基胞嘧啶代替。
2. 嘌呤碱基:两类核酸DNA和RNA都含有嘌呤碱基:腺嘌呤(A)和鸟嘌呤(G)。另外,自然界中还存在重要嘌呤衍生物,如次黄嘌呤、黄嘌呤、尿素酸等(图1.2)。除此之外,核酸中还存在一些含量甚少的碱基,称稀有碱基。稀有碱基的含量一般不超过碱基总量的5%,而大多数都是一些碱基的甲基化产物。

1.1.1.2 糖

RNA中,糖类是以呋喃型形式存在的D-戊糖(核糖, β -D-ribofuranose),而DNA则为D-2-脱氧戊糖(脱氧核糖, β -D-2-deoxyribofuranose, 图1.3)。

1.1.1.3 核苷

核苷是一种糖苷化合物,由戊糖和碱基缩合而成,糖与碱基之间以糖苷键相连接,戊糖的第一位碳原子(C_1)与嘧啶碱的第一位氮原子(N_1)或与嘌呤碱的第九位氮原子(N_9)相连接,构成N-C键(N-糖苷键,图1.4)。根据碱基种类的不同可形成不同的核苷:腺嘌呤核苷(adenosine)、鸟嘌呤核苷(guanosine)、胞嘧啶核苷(cytidine)、尿嘧啶核苷(uridine)、腺嘌呤脱氧核苷(deoxyadenosine)、鸟嘌呤脱氧核苷(deoxyguanosine)、胞嘧啶脱氧核苷(deoxycytidine)、胸腺嘧啶脱氧核苷(deoxythymidine)。

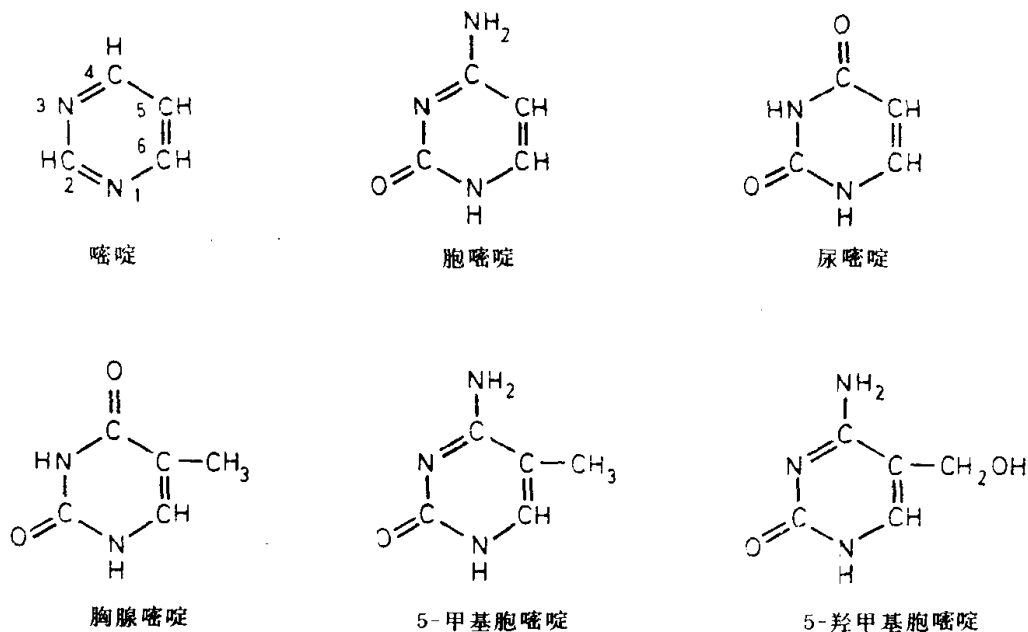


图 1.1 嘧啶碱基

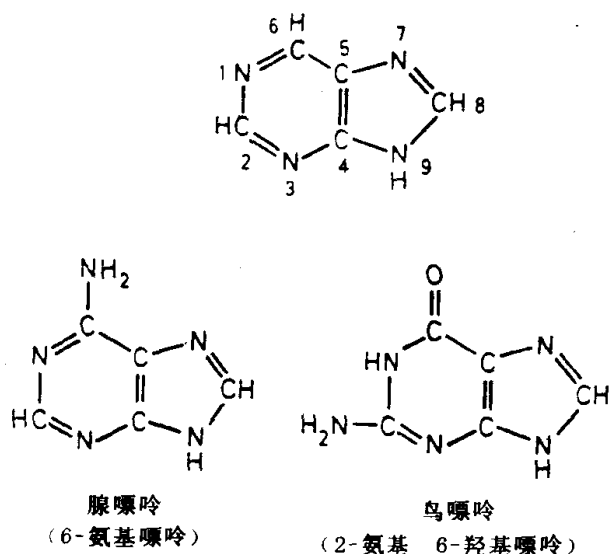


图 1.2 嘌呤碱基

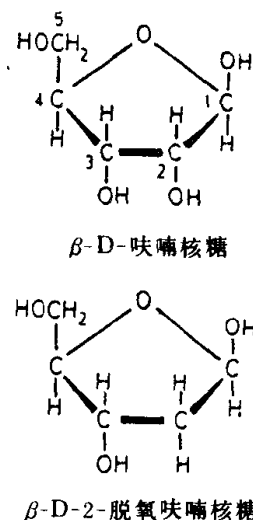


图 1.3 D-戊糖和 D-2-脱氧戊糖

1.1.1.4 核苷酸

核苷中的戊糖上的羟基被磷酸酯化所形成的化合物就是核苷酸。依照核苷酸中的戊糖的不同,可将核苷酸分成两大类:核糖核苷酸(ribonucleotides)与脱氧核糖核苷酸(deoxyribo-nucleotides)。由于核苷的糖环上有 3 个自由羟基,脱氧核苷的糖环上有 2 个自由羟基,则可分别形成 3 种和 2 种不同的核苷酸。生物体内游离存在的是 5'-核苷酸。5'-核苷酸的磷酸基团进一步磷酸化,则可形成 5'-二磷酸核苷酸(5'-nucleoside, diphosphate, 5'NDP)及 5'-三磷酸核苷酸(5'-nucleoside triphosphate, 5'-NTP),图 1.5 为腺苷酸、腺苷二磷酸、腺苷三磷酸结构图。

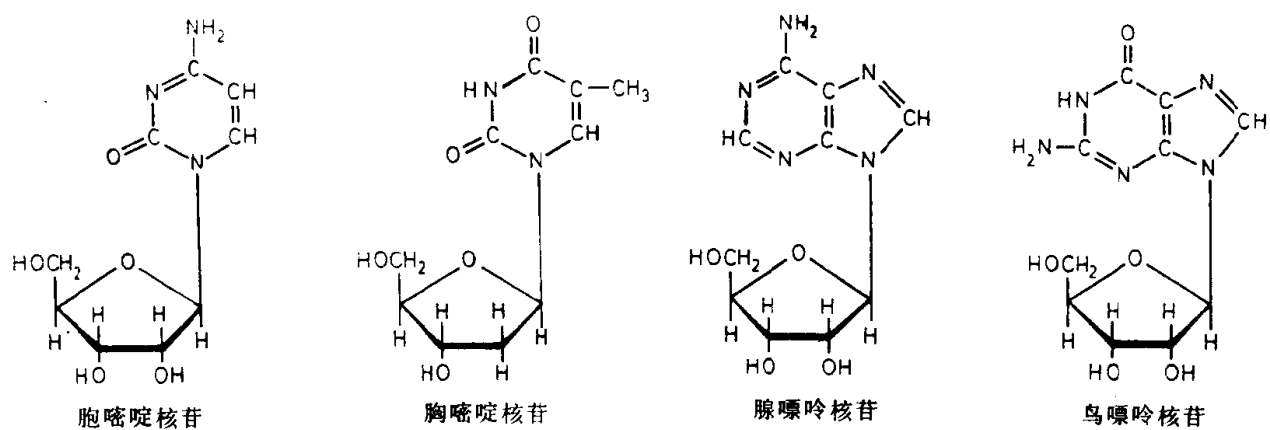


图 1.4 核苷

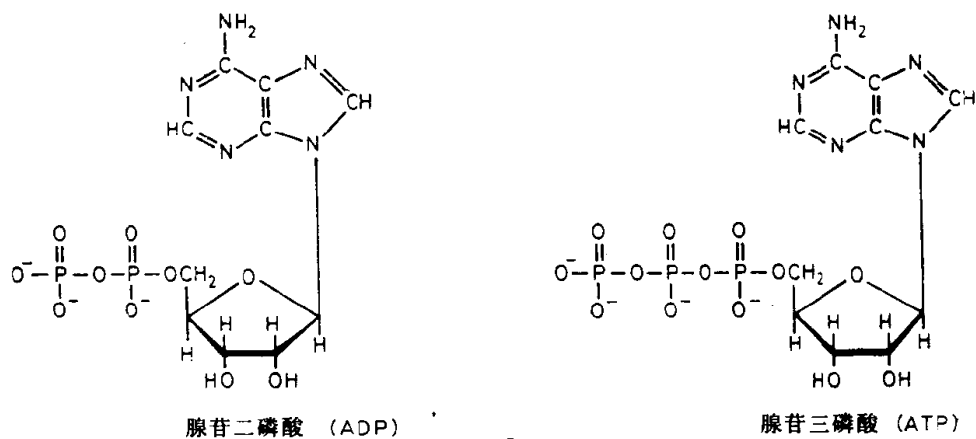
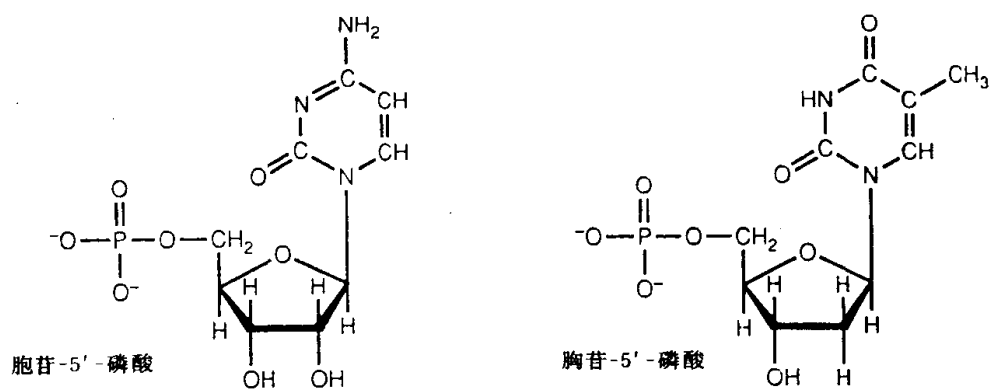
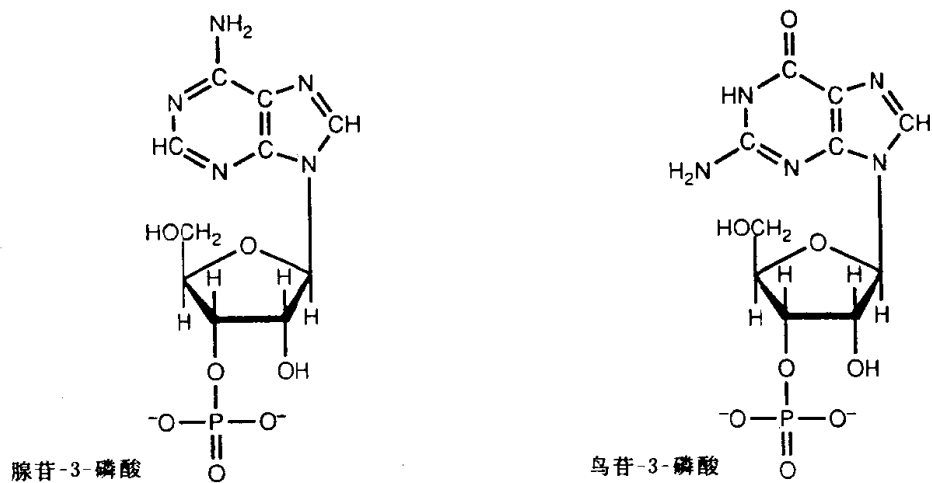


图 1.5 核苷酸

1.1.2 核酸的结构

核酸是由成百、成千乃至上百万个结构单位——核苷酸通过一定的连接方式构成的多体化合物。核酸的结构通常是指其一级结构及与其相对应的高级结构——空间结构。

1.1.2.1 核酸的一级结构

核酸的一级结构主要是指核苷酸之间的化合键及核苷酸的排列顺序。DNA 的一级结构是由 4 种脱氧核糖核苷酸通过 3',5'-磷酸二酯键彼此连接起来的直线形或环形分子。由于脱氧核糖中的第二位碳原子(C₂)上不含羟基,而第一位碳原子(C₁)又与碱基相连,因此,核苷酸之间的结合唯一可能形成的是 3',5'-磷酸二酯键。图 1.6 左示 DNA 的一级结构的局部。

RNA 的一级结构则是由核糖核苷酸[主要由 4 种核糖核苷酸(A,G,C,U)]组成,另外还存在一些稀有碱基组成的核苷酸组成的线状或环状多聚核苷酸,核苷酸之间的连键也为 3',5'-磷酸二酯键,不形成 2',5'-磷酸二酯键。图 1.6 右表示 RNA 的一级结构局部。

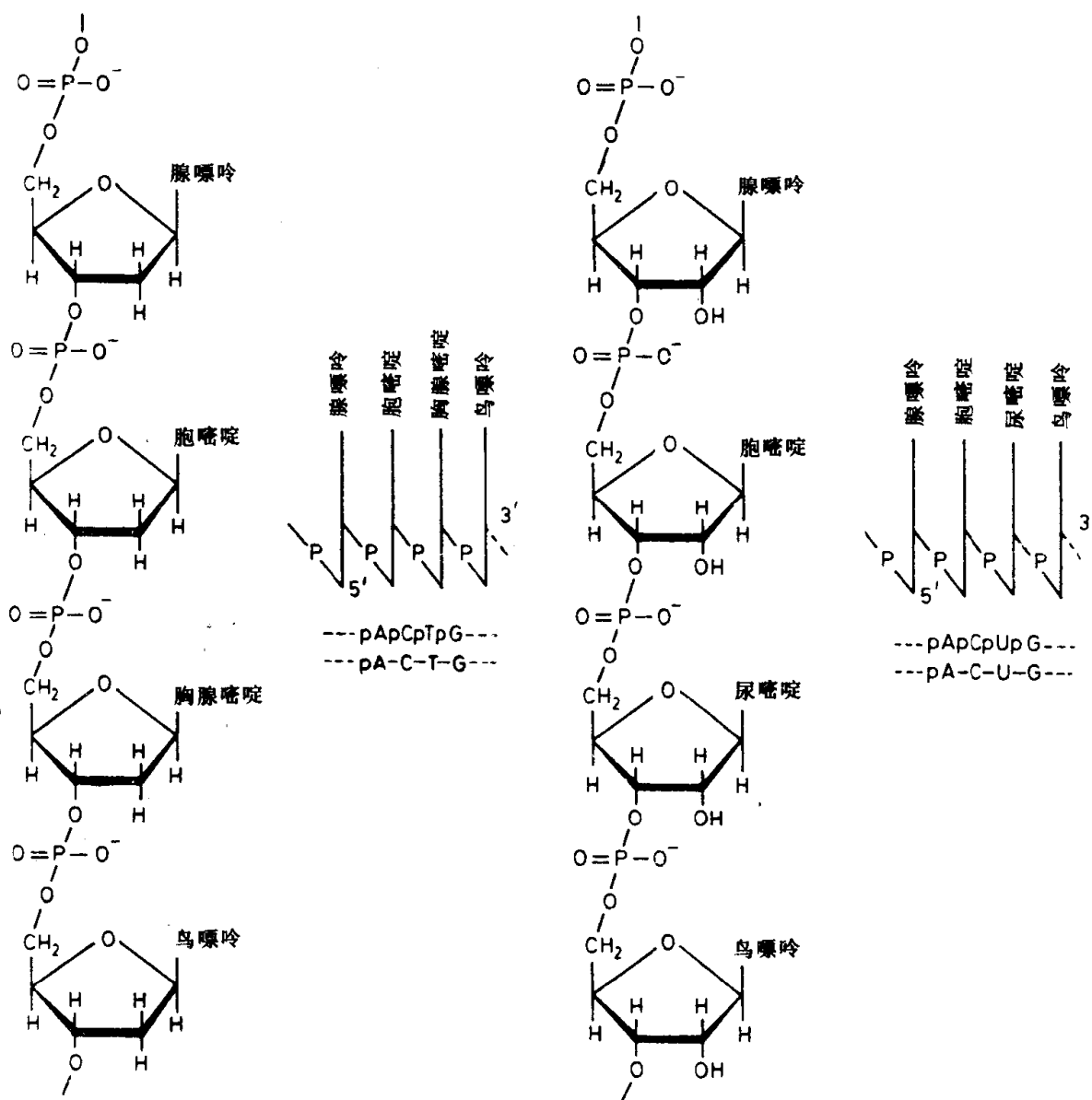


图 1.6 DNA(左)和 RNA(右)一级结构示意图。示意图右侧为缩写式。

在核酸的一级结构中核苷酸的排列顺序(序列)是分子生物学所要解决的一个重要课题。由于生命的遗传信息储存在核酸的序列之中(对于高等生物而言,DNA 是其遗传信息的载体,而对于一些病毒或更微小的生命体如类病毒而言,RNA 可能是其遗传信息的唯一载体),因此如果能解决各种生物的核酸序列,则可更深入地了解生命活动的本质所在。目前方法学上的发展提供了解决这一问题的可能性。实际上,人类已着手对包括人类在内的多种生物进行全基因组的序列分析。本书将试图对核酸序列分析作一原理与方法学上的系统介绍,以使广大读者更好地掌握这一技术而有利于进行分子生物学的深入研究。

1.1.2.2 DNA 的二级结构及三级结构

较早的研究表明,所有 DNA 中的腺嘌呤与胸腺嘧啶的摩尔含量相等,即 $A=T$;鸟嘌呤与胞嘧啶(包括 5-甲基胞嘧啶)的摩尔含量相等,即 $G=C$ 。因而,嘌呤的总数等于嘧啶的总数,也即 $A+G=C+T$ 。所有 DNA 中 $A=T, G=C$ 这一规律的发现,为 DNA 双螺旋结构模型的建立提供了重要的依据。

Watson 和 Crick 两人所建立的 DNA 双螺旋结构模型是目前所公认的 DNA 二级结构,DNA 由二条反向平行的多核苷酸链绕同一中心轴构成双螺旋结构。磷酸基团与戊糖构成的链条在外侧,碱基平面与轴垂直,位于内侧。碱基间的堆积距离为 0.34nm,两个核苷酸间的夹角为 36° ,每匝螺旋由 10 对碱基组成。DNA 的直径为 20nm。碱基之间的配合必定是 $A-T, G-C$ 。A 与 T 配对,其间形成 2 个氢键;G 与 C 配对其间形成 3 个氢键,因此,G 与 C 的配对更为稳定(图 1.7)。

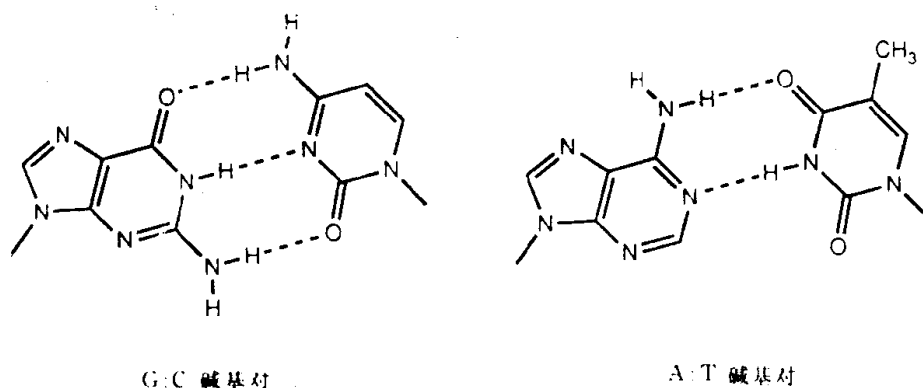


图 1.7 DNA 中碱基配对的一般原则

DNA 双螺旋结构非常稳定,其原因是有多种力量维持:氢键、碱基堆积力以及磷酸残基上的负电荷与介质中的阳离子之间形成的离子键。

DNA 双螺旋结构模型的提出,阐明了碱基配对的特异性,也即刻提供了 DNA 复制的机制,正如 Watson 和 Crick 自己所阐述的“如果两条链中的一条上的碱基顺序是给定的,那么人们就可以写出另一条链的碱基的准确顺序,因为它们之间有特异的配对关系。因此一条链就是另一条的互补链”,两条链“实际上是一对模板,其中的一个与另一个互补。我们想象在复制之前氢键断裂,两条链松展而分开,然后每条链都可作为模板为其本身形成一条伙伴链……碱基对的顺序也被精确地复制出来了。”

在二级结构(双螺旋结构)的基础上,DNA 还可通过折叠卷曲形成三级结构,双链环型 DNA(double-stranded cyclic DNA)的超螺旋(superhelical form)和开环型(open cir-

cular form),例如细菌质粒 DNA(plasmid DNA)、噬菌体 DNA 均存在超螺旋及开环型两种不同的形式。而目前有证据表明,自然界还存在着三链螺旋结构 DNA。

1.1.2.3 RNA 的高级结构

在细胞内 RNA 分子具有多种多样的功能。不同类型的 RNA 分子,其 RNA 链长度、二级结构及三级结构均不相同。不论是动物、植物还是微生物,其细胞内都含有 3 种主要的 RNA,即核糖体 RNA(ribosomal RNA,rRNA)、转运 RNA(transfer RNA,tRNA)、信使 RNA(messenger RNA,mRNA)。RNA 为单链分子,但它通过链内的折叠使碱基互补的区域发生碱基配对,形成氢键,同时形成局部双螺旋结构。而不能配对的区域形成突环(loop)。RNA 中双螺旋结构的稳定因素主要是碱基堆积力,其次为氢键,每一段双链区至少需要 4—6 对碱基,才能保持稳定。不同的 RNA 分子中,双链所占的比例不同,如 rRNA 中占 40%,tRNA 占 50%,而类病毒 RNA 则更高。

1.2 核酸的变性与复性

1.2.1 核酸的变性(denaturation)

由于高温或极端 pH,导致核酸分子双螺旋的氢键断裂,双链分子不再结合在一起,这就是所谓的核酸变性。

引起核酸变性的因素很多。温度升高,酸碱度的改变,乙醇、丙酮、尿素、酰胺等有机物质都可使核酸发生变性。因温度升高而引起的变性称热变性,而因 pH 的改变引起的变性称作酸碱变性。

当 DNA 因介质环境改变而发生变性如热变性时,其双链间的氢键断裂,双螺旋结构发生破坏,两条链彼此分离,形成无规则线团。随着 DNA 结构上的改变,光吸收增加(260nm 处紫外吸收值升高),比旋下降,粘度降低,酸碱性改变。在某一温度下,变性的 DNA 分子占 DNA 总量的 50%,这一温度称为这种 DNA 的溶解温度,以 T_m 表示。DNA 的 T_m 值与 DNA 的 GC 含量、溶剂的离子强度以及 DNA 的均一性有关。GC 含量越高, T_m 值越高;浓度越高,则同种 DNA 的 T_m 值也越高。

RNA 的变性也导致无规则线团的形成。但是因为 RNA 只有局部的双螺旋区,所以这种转变不如 DNA 那样明显, T_m 值也较低。

1.2.2 核酸的复性(renaturation)

变性 DNA 在适当的条件下,两条分离的单链可以重新发生碱基配对,恢复双螺旋结构,这一过程即所谓的复性,复性后的 DNA 可得到物理化学性质上的恢复。复性的条件与 DNA 样品的性质、介质温度的变化以及介质的离子强度均有关系。

复性在分子生物学研究中有着相当广泛的应用。核酸分子杂交、DNA 合成反应中引物与模板 DNA 的退火(annealing),均是核酸复性的应用。

1.3 核酸酶与核酸的降解

几乎在所有的生物体系中都存在着能够水解磷酸二酯键使核酸断裂的酶类。一些酶

类专一性地作用于 RNA,称作核糖核酸酶(ribonucleases, RNase); 另外一些酶类称作脱氧核糖核酸酶(deoxyribonucleases, DNase), 这类酶专一性地针对 DNA; 还有一类酶非特异性地作用于 RNA 与 DNA。此外, 还存在磷酸化酶类, 如多核苷酸磷酸化酶与焦磷酸化酶能够使 RNA 解聚, 但其体内的降解作用不清楚。磷酸单酯酶作用于带有磷酸基团末端的多聚核苷酸或寡核苷酸以及其他单核苷酸以释放出无机磷酸, 它们作用的底物常常是核酸酶作用的产物。

核酸酶种类很多, 根据不同的分类标准可将核酸酶类分成不同的类型。按对底物作用专一性分为核糖核酸酶与脱氧核糖核酸酶; 按对底物作用的方式可分为核酸内切酶(endonuclease)与核酸外切酶(exonuclease)等。

核酸酶类以及其他的一些核酸修饰酶, 是分子生物学家在进行研究及核酸操作中不可缺少的工具酶。有了这些酶类, 当今生物领域中的基因工程及核酸序列分析才得以迅速发展。

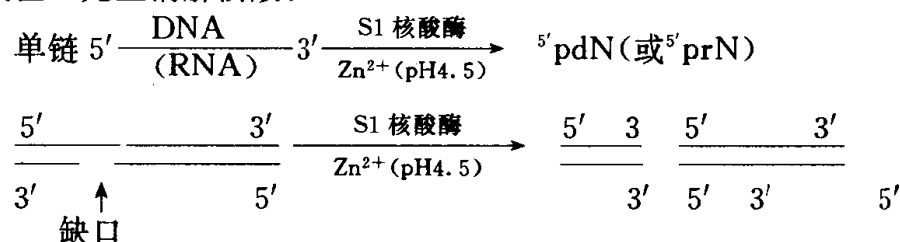
1.3.1 非专一性核酸酶

非专一性核酸酶的作用底物为 DNA 和 RNA, 一般要求 DNA 为变性分子或单链分子。

1.3.1.1 非专一性内切酶

这类酶的作用位点在多核苷酸链的内侧, 可对某种核苷酸有专一性。分解产物或是以 5'-磷酸为末端的核苷酸或寡核苷酸的混合物, 或是 3'-磷酸为末端的核苷酸或寡核苷酸的混合物。这些酶的活性要求某种二价金属离子, 如 Mg^{2+} 或 Zn^{2+} 或 Ca^{2+} 等。小球菌核酸酶(micrococcal nuclease)、粗糙链孢霉核酸酶(neurospora crassa nuclease)、S1 核酸酶(S1 nuclease)、绿豆核酸酶(mung bean nuclease)等均属非专一性内切酶。

1. S1 核酸酶: S1 核酸酶的作用底物为单链 DNA 或 RNA, 分解产物是以 5'-磷酸为末端的单核苷酸或寡核苷酸。双链 DNA 及 RNA 或者 DNA: RNA 杂交分子对该酶相对不敏感。然而如果酶达到一定量或相当过量时, 该酶可在切口处或缺口切割双链核酸甚至完全消解核酸。



S1 核酸酶可有以下三个方面的应用: ① 分析 DNA: RNA 杂交分子的结构; ② 消解 DNA 片段中突出的单链部分以产生平端分子; ③ 打开 cDNA 分子中的发夹环结构。

2. 绿豆核酸酶: 绿豆核酸酶的作用底物与产物与 S1 核酸酶相同, 但绿豆核酸酶的作用强度比 S1 核酸酶要温和一些。绿豆核酸酶的主要用途是消解 DNA 分子中突出的单链部分以产生平端分子。

1.3.1.2 非专一性外切酶

这类酶的作用点从多核苷酸的末端开始, 逐个地将核苷酸切割下来。切割产物或是以

5'-磷酸为末端的核苷酸,或是以 3'-磷酸为末端的核苷酸。蛇毒磷酸二酯酶(venom phosphodiesterase)以及牛脾磷酸二酯酶(spleen phosphodiesterase)即属该类酶,前者从 3'-OH 末端开始切割(有 3'-磷酸末端的寡核苷酸不被作用),产物为 5'-磷酸末端核苷酸;后者从 5'-OH 末端开始水解,产物为 3'-磷酸末端核苷酸。

1.3.2 核糖核酸酶

核糖核酸酶为专一性核酸酶,其作用底物为 RNA。也分为内切酶与外切酶。这些酶类在实验室被广泛地应用于切割与消解 RNA,如用它们进行 RNA 的序列分析、消除 DNA 样品中的 RNA 分子等。表 1.1 罗列了这类酶的来源、类型、产物及功能。

表 1.1 核糖核酸酶

酶	来源	类型	产物	底物特异性	功能
胰 RNaseA	哺乳动物胰脏	内切酶	3'-PO ₄ , 5'-OH	PypN	降解
RNaseT1	<i>Aspergillus oryzae</i>	内切酶	3'-PO ₄ , 5'-OH	GpN	降解
RNaseT2	<i>Aspergillus oryzae</i>	内切酶	3'-PO ₄ , 5'-OH	NpN	降解
RNaseU1	<i>Ustilago sphaerogena</i>	内切酶	3'-PO ₄ , 5'-OH	GupN	降解
RNaseU2	<i>Ustilago sphaerogena</i>	内切酶	3'-PO ₄ , 5'-OH	PupN	降解
RNaseL	哺乳动物细胞	内切酶	3'-PO ₄ , 5'-OH	UypN	降解
RNase I	<i>E. coli</i>	内切酶	3'-PO ₄ , 5'-OH	非特异性	降解?
RNaseP1	<i>Penicillium citrinum</i>	内切酶	5'-PO ₄ , 3'-OH	ssRNA 或 DNA	?
绿豆核酸酶 I	绿豆芽	内切酶	5'-PO ₄ , 3'-OH	ssRNA 或 DNA	?
小球菌核酸酶	<i>Staphylococcus</i>	内切酶	3'-PO ₄ , 5'-OH	ssRNA 或 DNA	降解
S1 核酸酶	<i>Aspergillus</i>	内切酶	5'-PO ₄ , 3'-OH	ssRNA 或 DNA	降解
内切核酸酶 H	多来源	内切酶	5'-PO ₄ , 3'-OH	RNA; DNA 分子中的 RNA	降解
外切核酸酶 H	逆转病毒	外切酶 (5'→3' 和 3'→5')	5'-PO ₄ , 3'-OH	RNA; DNA 分子中的 RNA	降解
RNase I	<i>E. coli</i>	外切酶 (3'→5')	5'-NMP	非特异性 (ssRNA 更敏感)	降解
蛇毒磷酸二酯酶	蛇	外切酶 (3'→5')	5'-NMP	RNA 或 DNA (ssRNA 更敏感)	降解
牛脾磷酸二酯酶	牛脾	外切酶 (5'→3')	3'-NMP	RNA 和 DNA (ssRNA 更敏感)	降解
多核苷酸磷酸酶	微生物	外切酶 (3'→5')	5'-NDP	可逆反应, 非特异性	未知
RNase III	<i>E. coli</i>	内切酶	5'-PO ₄ , 3'-OH	dsRNA (不知识别机制)	加工
RNaseI V	<i>E. coli</i>	内切酶	3'-PO ₄ , 3'-OH	特异位点 (未知)	未知
RNase P	<i>E. coli</i>	内切酶	5'-PO ₄ , 3'-OH	tRNA 前体	加工
RNase E	<i>E. coli</i>	内切酶	5'-PO ₄ , 3'-OH	ssRNA 前体	加工
RNase M5	<i>E. coli</i>	内切酶	可能 5'-PO ₄ , 3'-OH	5S, 23S 及 16S 前体	加工
M16, M23					
RNase F, Q, Y, D, T, PH 和 BN	<i>E. coli</i>	内外切酶	可能 5'-PO ₄ , 3'-OH	tRNA 前体	加工

1.3.3 脱氧核糖核酸酶

1.3.3.1 外切酶

分别从原核生物和真核生物中分离出大量的脱氧核糖核酸外切酶。表 1.2 示来源于

大肠杆菌(*E. coli*)的一些外切酶。这些酶的作用底物、作用方式以及分解产物均不相同,在 DNA 分子操作中各有用途。

表 1.2 大肠杆菌核糖核酸外切酶

名 称	基 因	分子量	作用底物	底物末端基团	产 物
外切酶 I	<i>sbc B</i>	7 2000	单链 DNA	3'-OH	5'-磷酸单核苷和双核苷
外切酶 II	<i>pol A</i>	109 000	单链或缺口双链 DNA	3'-OH	5'-磷酸核苷
	<i>pol B</i>	120 000	单链 DNA	3'-OH	5'-磷酸核苷
	<i>pol C</i>	140 000	单链或缺口双链 DNA	3'-OH	5'-磷酸核苷
外切酶 III	<i>xth A</i>	28 000	双链(AP 位点)	3'-OH 或 3'-OP	5'-磷酸核苷和单链 DNA
外切酶 IV	—	—	寡核苷酸	3'-CH	5'-磷酸核苷
外切酶 V	<i>rec B</i>	140 000	双链或	3'-OH 或	寡核苷酸
	<i>rec C</i>	130 000	单链 DNA	5'-OH	
外切酶 VI	<i>pol A</i>	109 000	缺口双链, RNA/DNA	5'-OH 或	短核苷酸
	<i>pol C</i>	140 000	杂交分子和单链 DNA	5'-OP	
外切酶 VII	<i>xse A</i>	88 000	单链 DNA	3'-OH 或 5'-OH	寡核苷酸
外切酶 VIII	<i>rec F</i>	(140 000) ₂	双链 DNA	5'-OH 或 5'-OH	5'-磷酸核苷

1.3.3.2 内切酶

首先纯化与定性的脱氧核糖核酸酶是两种内切酶:来源于牛胰脏的牛胰脱氧核糖核酸酶(pancreatic deoxyribonuclease, DNase I)以及来源于牛脾和胸腺的牛脾脱氧核糖核酸酶(spleen deoxyribonuclease, DNase I)。此后陆续从原核生物及真核生物中分离了许多 DNA 内切酶(表 1.3)。

表 1.3 大肠杆菌脱氧核糖核酸内切酶

名 称	基 因	分子量	作用底物	产 物
内切酶 I	<i>end A</i>	12 000	双链 DNA	5'-磷酸寡核苷酸
内切酶 II	—	未知		
内切酶 III	—	27 000	AP 位点和紫外辐射作用过的双链	带 3'-脱氧核糖和 5'-磷酸末端的缺口
内切酶 IV	—	33 000	AP 位点	带 3'-OH 和 5'-脱氧核糖末端的缺口
内切酶 V		27 000	单链 DNA 损伤双链	短寡核苷酸
内切酶 VI	<i>xth A</i>	28 000		
uvrABC	<i>uvr A</i>	114 000	损伤 DNA	短的切割区域
内切酶	<i>uvr B</i>	84 000		
	<i>uvr C</i>	68 000		
限制性内切酶	<i>hsd R</i>	135 000	非修饰识别系列的	长的双链片段
	<i>hsd S</i>	60 000	双链 DNA	
	<i>hsd M</i>	55 000		

1.3.3.3 限制性内切酶

限制性内切酶能识别特殊的双链 DNA 序列并在这一序列的一定位置上把 DNA 切断。目前这类酶主要发现于细菌。分为 3 种类型: I, II, III, 其中 I 类限制性内切酶除切割活性外还具有甲基化酶及 ATPase 的活性; II 类限制性内切酶除切割活性外还具有甲基化酶的活性, 因此, 这两类酶在分子克隆中都不常用。

第 III 类内切酶在分子克隆中应用最大。比较起第 I, II 类酶, 第 III 类酶则具有以下一些重要的特性: ① 有特定的识别序列, 通常为 4—6 个碱基对, 多者可达 8—10 个碱基对, 而且大多数具有回文对称结构和旋转对称轴; ② 切割位点位于识别序列的固定位置上,