



生命科学前沿及应用生物技术

基因克隆与操作

(原书第二版)

〔英〕克里斯托弗·豪 著

李慎涛 程 杉 主译

Gene Cloning and Manipulation

Second Edition



科学出版社

生命科学前沿及应用生物技术大系·典藏版

基因克隆与操作

(原书第二版)

Gene Cloning and Manipulation

Second Edition

[英] 克里斯托弗·豪 著

李慎涛 程 杉 主译

科学出版社

北 京

图字: 01-2009-0733 号

内 容 简 介

应用生物技术大系和现代生命科学前沿系列图书分别被列为“十一五”和“十二五”国家重点图书出版规划项目。本丛书针对生命科学领域前沿重点发展方向以及应用生物技术领域的新成果、新思路、新方法和新技术,全面展示了其最新的发展动态,涵盖了基础理论和主要技术方法,呈现了新的概念与理论、技术,在更深层次上阐明了生命的本质规律,给人们提供了新的认识生命本质的手段,也为生物技术服务于人类开辟了新的途径。涉及领域包括生物医药、干细胞技术、工业微生物学、蛋白质及蛋白质组、系统生物学、合成生物学、生物材料、农业生物技术、环境生物技术、海洋生物技术、生物资源与安全等。

Gene Cloning and Manipulation Second Edition 978-0-521-52105-5 by Christopher Howe first published by Cambridge University Press 2007

All rights reserved.

This simplified Chinese edition for the People's Republic of China is published by arrangement with the Press Syndicate of the University of Cambridge, Cambridge, United Kingdom.

© Cambridge University Press & Science Press 2010

This book is in copyright. No reproduction of any part may take place without the written permission of Cambridge University Press or Science Press.

This edition is for sale in the mainland of China only, excluding Hong Kong SAR, Macao SAR and Taiwan, and may not be bought for export therefrom.

此版本仅限中华人民共和国境内销售,不包括香港、澳门特别行政区及中国台湾。不得出口。

图书在版编目(CIP)数据

生命科学前沿及应用生物技术大系:典藏版/舒红兵等编著. —北京:科学出版社, 2016

ISBN 978-7-03-047487-2

I. ①现… II. ①舒… III. ①生命科学—研究②生物工程—研究 IV. ①Q1-0②Q819

中国版本图书馆 CIP 数据核字(2016)第 043876 号

责任编辑:王 静 李 悦

责任印制:张 伟 / 封面设计:刘新新

科学出版社 出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

北京厚诚则铭印刷科技有限公司 印刷

科学出版社发行 各地新华书店经销

*

2016 年 7 月第 一 版 开本:787×1092 1/16

2016 年 7 月第一次印刷 印张:2108

字数:49 985 000

定价:8900.00 元

(如有印装质量问题,我社负责调换)

本书参译人员

主译

李慎涛	首都医科大学
程 杉	首都医科大学

译校者 (以姓氏拼音排序)

李 京	首都师范大学
李小璐	北京协和医学院
刘晓娟	首都医科大学
阮 晶	首都医科大学
翟方丽	首都医科大学
郑少鹏	首都医科大学

译者序

近年来，基因工程技术已得到飞速的发展，已应用到生物科学的方方面面。需要掌握和学习基因工程技术的人越来越多，而一本好的教材可以达到事半功倍的学习效果。《基因克隆与操作》来源于剑桥大学克里斯托弗·豪教授给大学生上生物化学和分子生物学课的系列讲稿。他集 20 年的教学经验精心编写的这本书内容丰富，涵盖了基因工程相关的原理、技术和方法；语言简练、通俗易懂，即使是初学者也可以很容易掌握。

这本书第一章介绍了基因克隆与操作所需要的各种酶和相关技术，如各种转化方法、电泳和印迹等；第二章详细介绍了 PCR 的原理、方法和应用，通过这一章的学习，可以全面地掌握 PCR 技术；第三、第四章介绍了克隆的方法，对常用的载体和宿主以及各种克隆技术进行了简明扼要的介绍，可以使读者灵活掌握各种克隆技巧；第五、第六章介绍了 DNA 文库和 cDNA 文库的构建、筛选及应用；第七章介绍了 DNA 改造的技术，通过学习，可以随心所欲地对 DNA 序列进行改造，以满足不同的需要；第八章介绍了克隆化 DNA 的应用，重点介绍了蛋白质的表达，包括蛋白质纯化，可以使读者初步了解克隆技术的应用；第九章介绍了大肠杆菌以外的其他细菌、真菌、藻类、植物、昆虫和哺乳动物的基因操作技术，可以使我们对基因工程技术有一个更全面的认识；第十章介绍了一些应用实例，可以使读者综合运用这本书的各种技术，解决某个具体的问题，开阔思路。

总之，这本书是一本很好的基因工程技术的教材，适合于学习生物相关专业的学生和从事相关研究的研究人员使用。

在这本书翻译过程中得到了首都医科大学丁卫教授、科学出版社莫结胜编辑等的大力帮助，在此表示衷心的感谢！同时，欢迎广大读者对书中翻译存在的谬误给予指正，不胜感激！

李慎涛

2010 年 4 月 12 日于首都医科大学

第二版前言

自本书第一版出版以来，基因克隆和操作的领域已发生了很大的变化，我在第二版中进行的修订反映了这种进展。PCR 方法的应用已得到了极大的发展，在许多生物中都可以用到“组学”和反向遗传学的技术，一些成熟的领域也取得了很大的进步，如用于表达蛋白质的宿主和载体，以及使用荧光蛋白作为报告基因。像在第一版中那样，我已尽量强调有关我们所用载体的基本原理，并尽量避免使用长长的、详细的列表（从任何角度来看，这些列表都很快就会过时）。由于认识到了设计针对各种实验状态的适当策略的必要性，我增加了最后一章，给出了一些实例和建议。

感谢我实验室的成员，当为了完成此版本（我的众所周知的“兴趣巨著”）而不得不推迟其他工作时，他们耐心地等待。我要特别感谢那些以各种方式给予直接帮助的人，特别是 Mim Bower、Jon Burton、Ellen Nisbet、Saul Purton、Beatrix Schlarb-Ridley 和 Petrus de Vries。我也要感谢剑桥大学出版社的 Katrina Halliday 和 Clare Georgy，以及 Keyword 集团的 Peter Lewis 和 Rasika Mathur，感谢他们的技术专长、耐心和鼓励。

克里斯托弗·豪

第一版前言

本书来自于给大学生上生物化学和分子生物学以及医学课的一系列讲稿，我希望本书能对学习生物科学相关学科的人有所帮助。我已尽量将重点放在与学科有关的基本原理上，而不是事无巨细地去解释载体系统和实际操作，这方面有更详细的专著、综述、公司目录和实验室手册。我要感谢许多朋友和同事，他们通过阅读部分章节（在某些情况下是全篇）或以其他方式促使本书完成，特别感谢 Janet Allen、Alison Baker、Adrian Barbrook、Alison Franklin、Hilary、Tony Larkum 和 Saul Purton，还要感谢剑桥大学出版社的 Robin Smith 在本书写作过程中给予的建议和鼓励，以及 Bookworks 的 Robert Sugar 和 Dorothy Duncan 在本书出版过程中给予的帮助。

克里斯托弗·豪

目 录

译者序

第二版前言

第一版前言

第 1 章 工具.....	1
1.1 前言	1
1.2 切割	1
1.3 修饰	7
1.4 连接.....	10
1.5 转化.....	12
1.6 从大肠杆菌中纯化质粒 DNA	14
1.7 核酸的凝胶电泳.....	15
1.8 寡核苷酸合成.....	19
1.9 微阵列.....	20
第 2 章 聚合酶链反应	21
2.1 基本技术.....	21
2.2 预防措施和缺点.....	27
2.3 改良.....	30
第 3 章 简单克隆	36
3.1 基本实验.....	36
3.2 载体、转化和宿主.....	40
3.3 修饰.....	45
3.4 接头、衔接子和序列盒.....	48
第 4 章 用于大肠杆菌的其他载体系统	52
4.1 前言.....	52
4.2 BAC 载体	52
4.3 M13 噬菌体载体	53
4.4 λ 噬菌体	57
4.5 黏粒.....	65
4.6 Mu 噬菌体	65
4.7 P1 噬菌体	66
第 5 章 文库构建	69
5.1 前言.....	69
5.2 基因组文库.....	69
5.3 cDNA 文库.....	70

5.4 专业文库..... 75

第 6 章 文库筛选 81

6.1 前言 81

6.2 数据库筛选..... 81

6.3 编码功能的实验筛选..... 82

6.4 其他功能的筛选：报告基因..... 94

第 7 章 改造与诱变..... 100

7.1 前言 100

7.2 基于限制性内切核酸酶的方法 100

7.3 寡核苷酸定向诱变 102

7.4 选择正确的突变 107

7.5 灭活基因 108

第 8 章 克隆化 DNA 的应用 114

8.1 作为 DNA 使用 114

8.2 RNA 的合成..... 114

8.3 蛋白质的合成 115

8.4 体外翻译 116

8.5 体内表达 116

8.6 研究基因功能：报告基因和标签 125

第 9 章 使用其他生物..... 128

9.1 前言 128

9.2 细菌 128

9.3 酿酒酵母 133

9.4 其他真菌 139

9.5 藻类 141

9.6 维管植物 142

9.7 细胞器转化 149

9.8 秀丽隐杆线虫 153

9.9 昆虫 153

9.10 哺乳动物..... 158

第 10 章 实例 172

10.1 前言..... 172

参考文献..... 177

索引..... 184

第1章 工 具

1.1 前 言

对基因进行克隆和操作需要能对遗传物质（通常是 DNA，但有时是 RNA）进行切割、修饰和连接，并能检查所要操作分子参数（如大小）。假设我们知道所涉及物质（DNA、RNA 等）的结构，我们将首先介绍一下对其进行操作的现有工具，其中许多都是在细胞中具有重要生理功能的酶。为了理解为什么这些酶对于我们的工作是有用的，我们也应当知道其正常的功能。

对于某项特定的工作，选用哪种酶主要取决于以下两点。

(1) 纯化是否容易（即便宜）？这取决于它在细胞中的丰度以及将其与其他不希望的活性物质分开的难易程度。

(2) 它是否最适于此项工作？这取决于它的特异性（“准确性”）和比活（“速度”），以及它所催化反应的细节。

其他因素（如稳定性）也很重要。

用于基因操作本身的酶也可以使用基因操作技术来生产。使用克隆的基因能更容易地大量制备这些酶，并能够对基因进行改造，以“改善”其功能，也许可稍微改变这些基因所编码的酶的特性。

1.2 切 割

分解核酸的酶被称为核酸酶——分解 RNA 的酶被称为核糖核酸酶或 RNA 酶（RNase），分解 DNA 的酶被称为脱氧核糖核酸酶或 DNA 酶（DNase）。

分解线性核酸分子的方式有两种：从末端一点儿、一点儿地分解；或在分子内切割，将核酸分子分解成碎片。前者被称为外切核酸活性（希腊语 *exo* 意为外部的）；后者被称为内切核酸活性（希腊语 *endo* 意为内部的）。不要错误地认为内切核酸酶是从末端向内切割，它是在内部将核酸分子切割成碎片。使用最广泛的内切核酸酶是限制性内切核酸酶。

1.2.1 限制性内切核酸酶

限制性内切核酸酶是细菌抵抗外来 DNA 的天然防御机制的一部分，这种外来 DNA 可能来自于病毒或某种外来细胞群体的质粒。最初认识到这些酶是发现在特定的大肠杆菌菌株中，它们能够限制某些病毒的生长，因此而得名（现在许多分子生物学工作者广泛地使用动词“限制”来表示“用一种限制性内切核酸酶切割”）。限制性内切核酸酶与修饰酶（对 DNA 甲基化）相关，甲基化保护 DNA 免受内切核酸酶的切割，这终止了细胞降解其自身的 DNA。尚未被正确甲基化的入侵 DNA 会被降解，除非这种

入侵 DNA 能够足够快地被细胞的甲基化酶修饰，但这种情况极少发生。

一旦 DNA 被修饰，那么即使在复制后也仍然保持被保护状态，这是因为两条链都被甲基化的 DNA 分子半保留复制，生成半甲基化的（即在一条链上甲基化）子代分子，半甲基化足以抵抗内切核酸酶的切割。在复制再次发生之前，未甲基化的链可以被修饰。

已识别出三种类型的限制/修饰系统，称为 I、II 和 III 型（或类），其主要特点见表 1.1。所有的酶都识别特定的 DNA 序列，但只有 II 型内切核酸酶在这些识别序列内切割，许多 II 型内切核酸酶的识别位点见表 1.2。这些酶通常使分子产生一个“交错”的切口，使分子（尽管主要是双链）具有短的单链末端，这些短的单链末端被称为黏性末端。

表 1.1 限制和修饰系统的特性

项目	I 型	II 型	III 型
组成	多酶；R（内切核酸酶）、M（甲基化酶）和 S（特异性）亚基的复合物，如 R ₂ M ₂ S	单独的酶；内切核酸酶是一种同型二聚体，甲基化酶是一种单体	M 亚基提供特异性；其本身的功能是一种甲基化酶；当与 R 亚基构成异源二聚体时，其功能是甲基化酶-内切核酸酶
辅因子 ^a	Mg ²⁺ 、ATP、SAM（切割和甲基化时需要）	Mg ²⁺ 、SAM（只有甲基化时需要）	Mg ²⁺ 、ATP（切割时需要），SAM（甲基化时需要，促进切割）
识别位点	不对称的，可以由两部分组成，也可以是简并的，如 <i>EcoK</i> （AACN ₆ GTGC）	对称的，可以由两部分组成，也可以是简并的（表 1.2）	不对称的，不间断的，5~6 个核苷酸长，如 <i>EcoP15</i> -CAGCAG。切割时需要方向相反但不一定相邻的两个拷贝；甲基化时需要一个拷贝
切割	离识别位点的距离不定（100~1000 个核苷酸）	在识别位点内，除了那些在外部切割（距离一定）的 II 型内切核酸酶以外	离识别位点 25~27 个核苷酸处
已定性的系统数目	几个，分成几个家族，例如，K 包括 <i>EcoB</i> 、 <i>EcoD</i> 、 <i>EcoK</i> 和其他	几百个	少数几个

a. ATP，腺苷三磷酸；SAM，S-腺苷甲硫氨酸。

表 1.2 II 型限制性内切核酸酶识别序列实例^a

<i>Apal</i>	G G G C C ' C C ' C C G G G	<i>AhaIII</i>	T T T ' A A A A A A ' T T T
<i>BamHI</i>	G ' G A T C C C C T A G ' C	<i>BglII</i>	A ' G A T C T T C T A G ' G
<i>Bsp120I</i>	G ' G G C C C C C C G G ' C	<i>DpnI</i>	G A ' T C C T ' A G
<i>DraI</i>	T T T ' A A A A A A ' T T T	<i>EcoRI</i>	G ' A A T T C C T T A A ' G
<i>HincII</i>	G T P y ' P u A C C A P u ' P y T G	<i>HindIII</i>	A ' A G C T T T T C G A ' A
<i>HpaII</i>	C ' C G G G G C ' C	<i>MaeIII</i>	' G T N A C C A N T G '

<i>NotI</i>	GC'GGCC GC	<i>PvuII</i>	CAG'CTG
	CG CCGC'CG		GTC'GATC
<i>SaII</i>	G'TCGA C	<i>Sau3A</i>	'GATC
	C AGCT'G		CTAG'
<i>SphI</i>	G CATG'C	<i>TaqI</i>	T'CG A
	C'GTAC G		AGC'T
<i>XbaI</i>	T'CTAG A		
	A GATC'T		

a. I 代表切割位点；N 代表任核苷酸；Py 和 Pu 分别代表嘧啶核苷酸和嘌呤核苷酸。

根据酶的不同，5'端或3'端会保持单链状态，所产生的分子在5'端有一个磷酸基，在3'端有一个羟基。少数的II型内切核酸酶正好在其识别位点的外侧切割，例如，*MboII*在其识别位点-GAAGA-3'端7个核苷酸处切割。除了在简并序列内以外，其他的内切核酸酶则在其识别位点内切割，例如，*MamI*在-GATNN'NNATC-处切割，在这里N可以是任意核苷酸。然而，根据这两种类型酶的生物化学特性，它们仍然都可以被认为是II型内切核酸酶。除了刚才提到的特殊情况以外，用某种已知II型内切核酸酶切割所产生的所有DNA分子，其末端都会具有相同的序列，但用I型内切核酸酶和III型内切核酸酶则不会这样，因为它们在其识别位点以外切割。由于被II型内切核酸酶切割的分子通常具有相同的末端，所以这样的分子能够互相配对，而且正如我们将要看到的，这样的分子可以通过DNA连接酶共价连接。一些II型内切核酸酶产生整齐的而不是交错的切口，在相同的位置切割双链（见表1.2），这样在分子上产生双链末端或平端也没有问题，因为用连接酶也可以连接平端分子。II型内切核酸酶的以下切割特性也是很重要的。

1. 识别位点通常在两条链上读起来是相同的（只要以相同的方向读，如5'→3'）。这种序列通常被称为回文结构。对用同一种酶切割的所有分子来说，退火并不需要识别序列是回文结构，尽管回文结构确实能够增加重新组合为所需构型的数目。例如，图1.1说明了用*HindIII*（识别序列为-AAGCTT-）切割的两个分子是怎样退火的——右侧分子的每一端与左侧分子的每一端退火。用识别序列不是回文结构的一种酶切割的两个分子也能够退火，但退火的可能性要小一些。
2. 大多数酶的识别位点为4或6个核苷酸。如果所有核苷酸出现的频率相等（在要切割的DNA和酶的识别位点中），而且是随机的，则平均每4⁴（即256）个核苷酸就会出现一个特定的4个核苷酸的基序（motif），所以，用带有这种位点的酶切割所产生的片段的平均长度为256bp。同样，使用识别序列为6个核苷酸的酶切割，所产生的片段的平均长度为4⁶（即4096）bp。实际上，情况并非如此，理由如下。
 - a. 在识别位点内，碱基出现的频率不相等。
 - b. 在要切割的DNA内，碱基出现的频率不相等，而且在整个基因组内，频率是可变的。
 - c. 碱基并不是随机出现的，例如，某些二核苷酸容易出现，而另外一些则不容易出现。在整个基因组的范围内，这种不随机性的程度通常是可变的。

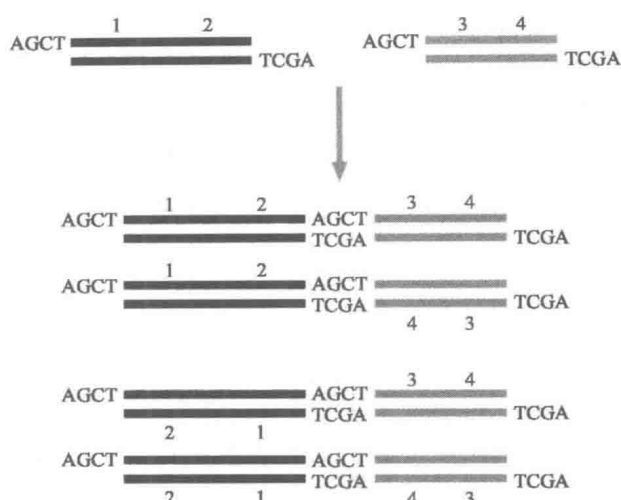


图 1.1 用 *Hind*III 切割的两个分子的退火。1、2、3 和 4 表示分子上的任意点，注意，由于 *Hind*III 切割位点为回文结构的特性，右侧分子能够以两种可能的方向与左侧分子的每一端退火。

3. 不同的酶能够识别相同的序列。例如，*Dra*I 和 *Aha*III 都识别和切割- $\text{TTT}'\text{AAA}-$ ，它们被称为同切点酶（isoschizomer）。然而识别序列相同的酶并不一定在识别序列内的相同位置切割，例如，*Apa*I 识别并切割- $\text{GGGCC}'\text{C}-$ ，而 *Bsp*120I 识别并切割- $\text{G}'\text{GGCCC}-$ 。
4. 不同的酶能够生成相同的末端。例如，*Sau*3AI 生成的末端为 $\text{GATC}-$ ，*Bam*HI 也生成同样的末端，这意味着用 *Sau*3AI 消化所生成的分子能够与用 *Bam*HI 消化所生成的分子退火和连接。假若平端分子也能够被连接，那么用任何产生平端的酶切割的分子都是可以相容的。然而，用不同酶切割的分子连接后可能不会再生成原有的识别位点。黏性末端相近但不完全互补的分子也可能能连接（尽管效率较低）。
5. 其他因素可以影响切割。在影响切割的因素中最重要的是：①甲基化；②所用的缓冲液；③底物的二级结构。

DNA 中碱基的甲基化可能来自于限制-修饰系统的修饰活性，也可能来自于许多独立甲基化酶的活性，常见的甲基化碱基包括 N^6 -甲基腺嘌呤、5-甲基胞嘧啶、5-羟甲基胞嘧啶和 N^4 -甲基胞嘧啶。限制性内切核酸酶通常不切割那些识别序列内特定碱基被甲基化了的分子，但在识别位点内，某些位置的甲基化可能不影响切割，而且对于某些酶，实际切割时可能需要其他位置的甲基化。例如，- GGATCC -识别位点内侧 C 的甲基化会抑制 *Bam*HI 的切割，但另一个 C 或 A 的甲基化则不会抑制；*Apy*I 识别位点（- CCAGG -或- CCTGG -）第一个 C 的甲基化抑制 *Apy*I 的切割，但其切割却需要第二个 C 的甲基化。

所用缓冲液会影响一些酶的特异性，例如，在正常情况下，*Eco*RI 在序列- GAATTC -处切割，但在浓度大于 5%（V/V）的甘油中，其特异性不严格，可能在- AATT -或- PuPuATPyPy -处切割，这种现象通常被称为星活性（star activity），

表示为 *EcoRI*。一些化合物能够改变切割的程度，如图 1.2 所示的溴化乙锭，这种分子可用于观看凝胶中的核酸，它能够嵌入（插入）到双链 DNA 分子的碱基之间，干扰限制性内切核酸酶的作用，使其只在一条链上切割。

在同一分子内，一些位点的切割效率要比其他位点低很多，这可能是由于 DNA 内的二级结构干扰内切核酸酶的识别或切割。

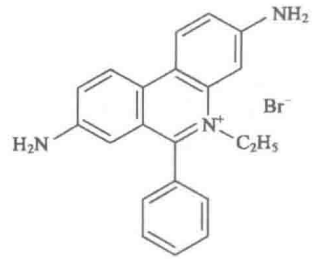


图 1.2 溴化乙锭。

6. 以单位来测量酶的活性。酶单位定义为在特定的条件（温度、pH 等）下，一定时间（通常为 1h）内消化标准量（通常是 $1\mu\text{g}$ ）标准类型的 DNA（通常是 λ 噬菌体或一种指定的质粒）所需要的酶量。所以，消化含有许多位点的 DNA 分子要比消化同样质量的含有较少位点的 DNA 需要更多单位的酶。
7. 用于克隆的限制性内切核酸酶制剂必须不含其他核酸酶。如果用于克隆的限制性内切核酸酶不纯，所生成的分子末端会被外切核酸酶降解，从而不能退火。污染的内切核酸酶活性会将分子切成没有（或错误）单链末端的片段，这也会引起同样的问题。所以，生产商通常用以下方法来检验酶制剂：将 DNA 与过量的酶温育，检测有多少比例的酶切产物能够被重新连接以及重新连接后的分子是否还能用酶切割。重新连接正确的比例越高，酶制剂则越“纯”。通过与末端被放射性标记的 DNA 分子的温育也可以检验酶制剂是否含有外切核酸酶，即底物 DNA 中释放出放射性标记表明含有外切核酸酶。然而，在简单的限制性内切核酸酶图谱分析（restriction enzyme mapping）中，低水平的核酸酶污染没有大碍。
8. 部分消化会很有用。有时我们故意用一种限制性内切核酸酶不完全消化，例如，我们可能会需要将一种生物制备的总 DNA（通常称为基因组 DNA）制成大小大致相同的片段（即 10kb），以便 10kb 片段库内含有生物的每一种序列。仅仅用识别序列为 6 个核苷酸的一种酶完全切割并收集所生成的大小约为 10kb 的片段是不适合的，许多 DNA 会被切割成太小（或太大）的片段，不会包括在 10kb 大小片段的库中。所以，更好的方法是使用一种切割很频繁（如在 4 个核苷酸识别位点处切割）的酶，但要调整允许的反应时间或反应中酶与 DNA 的比率，以便只有少数可能的位点被切割。这样就能够将片段的平均大小控制在 10kb 或任何其他所需要的大小，而且如果所有的位点被切割的机会或多或少地相等，那么在 10kb 的片段中会包含 DNA 的所有区域（除非在某个特定区域中位点分布很不正常）。在构建基因组文库时，这种方法很有用（见第 5 章）。
9. 使用标准的分子可以确定切割位点。用酶消化一些精选的、序列已完全知道的 DNA 分子，并将受试酶与其他切割位点已知的酶组合起来消化样品，用凝胶电泳检测所产生片段的大小，通过计算机分析，从序列能够推测出可能的识别位点（这些识别位点就是产生所观察到大小片段的位点），更准确地测量消化所产生分子的大小（至确切的核苷酸数目）能够推测出识别序列内实际的切割位点。也可以用电泳的方法准确测量这些片段的大小，用 DNA 测序反应的产物作为标志物。

10. 命名法遵循一种简单的约定。一旦一种酶被定性，则必须给其一个名称。约定名称以三个字母（斜体）开始：来源细胞的属名的第一个字母和种名的头两个字母。在适当的情况下，后面接着表明品系，然后是一个数字（罗马数字）表明来自名称所提到的那种品系的哪一种酶。例如，*EcoRI* 和 *EcoRII* 是指从大肠埃希菌（*E. coli*）R 菌株分离的第一种和第二种活性酶。名称通常被缩写，所以 *EcoRI* 通常只念成“RI”（发音为“R-one”），*BamHI* 通常只念成“Bam”，等等。根据酶命名法的通用约定，I 型、II 型和 III 型限制性内切核酸酶被分类为产生 5'-磷酸单酯的“脱氧核糖内切核酸酶”，并被分别分类为 EC 3.1.21.3、EC 3.1.21.4 和 EC 3.1.21.5。
11. 其他酶能够在特定的序列切割 DNA 分子。在 λ 噬菌体感染大肠杆菌细胞的过程中，噬菌体基因组的拷贝在一种特定的 16 个核苷酸位点处被切割，叫做 *cos*，形成一种 12 个核苷酸单链的突出，这种切割由一种被称为末端酶（terminase）的噬菌体酶来执行。有时将 *cos* 位点引入到大的 DNA 分子内，以便能够在单一的位点切割。当用常规限制性内切核酸酶切割时，能够将分子切成几个片段。

1.2.2 DNA 酶

在某些情况下，不适合用限制性内切核酸酶切割 DNA，如当 DNA 的碱基组成很不正常时。不过现在有这么多的酶可用，有很多的识别位点，故很难遇到这样的问题。一个较常见的问题是需要将 DNA 分解成平均大小仅为几百碱基对的随机片段库。用识别 4 个核苷酸的酶进行部分消化并不适合，因为要得到所需平均大小的片段，几乎每个位点都必须被切割，这将意味着一些序列仅在比所需大小范围小许多或大许多的片段中出现。使用 DNA 酶（如 DNA 酶 I）能够避免这个问题，因为 DNA 酶具有非常小的（对此目的来说，其实是没有）序列特异性，同样，仔细调整酶与 DNA 的比率或温育时间，可确保片段大小的最佳分布。使用 DNA 酶的一个问题是所生成的分子末端没有独特的单链序列，而且并非所有的末端都是平端，这使片段的克隆较为困难，但是用一种适当的 DNA 聚合酶将所有的末端补平则可以解决这个问题。

1.2.3 物理应力

除了酶学方法外，我们还可以使用物理剪切的方法随机切割 DNA。完成这种工作有几种方法，例如，简单地搅动溶液，或强迫溶液通过一个狭窄的孔（如注射针头或移液器头），或使用超声（提供高频振动）。在实际工作中，超声是最好的方法，因为它最容易控制，而且通常要比 DNA 酶 I 处理的重复性好。已有不同种类的超声破碎仪，最简单的是将一个金属探头浸入溶液中，使探头高频率地振动。这种方法的缺点是探头可能成为 DNA 制备物之间交叉污染的原因，除非仔细地清洗探头。还有一种杯状破碎头（cup-horn）超声破碎仪，这种仪器将所要破碎的溶液截留在—根漂浮在少量水中的管子内，探头浸入周围的水中，振动从水传播到含有样品的管子中。使用剪切的方法与使用 DNA 酶处理一样，无法控制所生成片段末端的序列。

1.3 修 饰

1.3.1 磷 酸 酶

磷酸酶是用水解的方法除去 DNA 分子上的磷酸基团而将其替换成羟基的酶。大多数的连接反应都需要限制性内切核酸酶切割后留下的末端磷酸基团，用磷酸酶封闭不要的连接反应将在第 3 章中介绍。最常用的酶制剂来自于小牛肠、河虾 (shrimp) 和南极嗜冷细菌 TAB5。许多磷酸酶制剂通过加热很容易被灭活，特别是最后一种，这在我们连接之前要终止磷酸酶活性时很有用。

1.3.2 聚 合 酶

我们将会遇到 4 种类型的 DNA 或 RNA 聚合酶：依赖 DNA 的 DNA 聚合酶、依赖 RNA 的 DNA 聚合酶、依赖 DNA 的 RNA 聚合酶和不依赖模板的聚合酶 (template-independent polymerase)。还有依赖 RNA 的 RNA 聚合酶，但这些酶对于我们的工作不太重要。

1. 依赖 DNA 的 DNA 聚合酶。这些酶使用 DNA 模板，以 $5' \rightarrow 3'$ 的方向合成一条 DNA 链，它们也可能具有 $5' \rightarrow 3'$ 和 $3' \rightarrow 5'$ 外切核酸酶活性，所有这些活性都可以用各种方式加以利用。所用的酶制剂来源于细菌 [如大肠杆菌和嗜热水生菌 (*Thermus aquaticus*)] 和用病毒 (如 T4 和 T7) 感染的细菌。广泛使用的大肠杆菌酶是 DNA 聚合酶 I，在正常情况下，它在 DNA 修复和用 DNA 引物替换 DNA 合成所需的 RNA 引物中具有重要作用。此酶具有 $5' \rightarrow 3'$ 聚合酶活性、 $3' \rightarrow 5'$ 外切核酸酶活性 (作为校对功能) 和 $5' \rightarrow 3'$ 外切核酸酶活性。其实这些活性位于分子的不同亚基上，用枯草芽孢杆菌蛋白酶切割可产生一个含有 $5' \rightarrow 3'$ 外切核酸酶活性的 35kDa 的 N 端片段和一个具有聚合酶活性及 $3' \rightarrow 5'$ 外切核酸酶活性的 76kDa 的 C 端片段 (图 1.3)。76kDa 的片段有时也被称为 Klenow 片段，完整的分子有时被称为 Kornberg 酶。

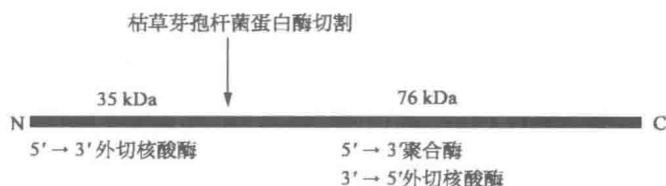


图 1.3 DNA 聚合酶 I。标明了活性的位置和枯草芽孢杆菌蛋白酶的切割位点。

$5' \rightarrow 3'$ DNA 聚合酶活性能用一种合适的模板合成一条互补的 DNA 链，这种模板可以是与一条小引物退火的大单链 DNA 片段，也可以是 $3'$ 端凹陷 (即 $5'$ 突出) 的限制酶切片段。将这些模板的任意一种与 DNA 聚合酶及正确的脱氧核苷三磷酸温育会使单链区 (凹陷末端) 充填，生成平端分子，这通常被称为末端补平 (end filling)。注意，Kornberg 酶的 $5' \rightarrow 3'$ 核酸外切活性通过降解突出的 $5'$ 端 (而不是通过合成其互补链) 也能够生成这样一个分子。通过聚合酶活性不能将凹陷的 $5'$ 端变成平端，因为合成必

须是 3'→5' 的方向，而这是不可能的。在这种情况下，3'→5' 核酸外切活性能够通过降解突出的 3' 端将末端变成平端。用任何方法将突出的末端变成平端都被称为修整 (polishing)，见图 1.4。大肠杆菌 DNA 聚合酶也被用于 DNA 测序。

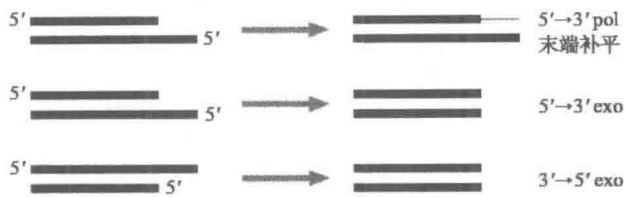


图 1.4 使用聚合酶 (pol) 或外切核酸酶 (exo) 活性修整突出的末端。细线表示通过聚合酶活性新合成的 DNA。

用聚合酶链反应 (PCR) 扩增 DNA 时，热稳定性 DNA 聚合酶是特别重要的，如在第 2 章中所叙述的那样。这些酶是从极端嗜热菌 (通常生长在高温的海底口) 分离而来的，如嗜热水生菌 (*Thermus aquaticus*)、嗜热高温球菌 (*Thermococcus litoralis*) 和强烈炽热球菌 (*Pyrococcus furiosus*)，其中一些菌生长在 100℃ 以上的温度中，在这种条件下，其 DNA 聚合酶也能够有效地发挥功能。在 PCR 中最常用的一种酶来自于嗜热水生菌，最适温度为 75~80℃，有关 PCR 所用酶的更多细节见第 2 章。热稳定性聚合酶还有其他用途，如用于 DNA 测序。

病毒 DNA 聚合酶由病毒基因组编码，一旦其感染大肠杆菌细胞便可合成 DNA 聚合酶。在没有宿主 DNA 合成时，病毒 DNA 的复制需要聚合酶。T4 DNA 聚合酶具有特别强的 3'→5' 外切核酸酶功能 (比 Klenow 酶的活性高 200 倍)，这在修整黏性末端时特别有用。T7 DNA 聚合酶有时也被用于 DNA 测序，且测序时比大肠杆菌 DNA 聚合酶 Klenow 片段产生的错误少，这可能是由于它具有持续的合成能力 (即不容易从模板上解离)。T7 DNA 聚合酶没有 5'→3' 外切核酸酶活性，用于测序的酶制剂 3'→5' 外切核酸酶活性通常被灭活，这种灭活最初是通过化学的方法来实现的，后来则是通过改变聚合酶的基因从而生成一种被称为“测序酶”的酶来实现。

2. 依赖 RNA 的 DNA 聚合酶。这些酶也被称为反转录酶，通常缩写为 RTase，因为它能够反转“中心法则”中转录信息的通常流向。反转录酶由反转录病毒编码 (在 *pol* 基因中)。反转录病毒有一种 RNA 基因组，在其部分生命周期中，这种 RNA 基因组必须转变为 DNA (之后 DNA 被插入到宿主的基因组中)。现广泛使用来自于鸟类成髓细胞白血病病毒 (avian myeloblastosis virus, AMV) 和莫洛尼鼠白血病病毒 (murine Moloney leukaemia virus, M-MLV) 的酶，修饰后的反转录酶 (如热稳定性提高的酶) 也可得到。与其他 DNA 聚合酶一样，反转录酶需要一个通过氢键与模板结合的引物，从而指导 DNA 以 5'→3' 的方向合成。在病毒感染的细胞中，引物是一种与部分反转录病毒基因组互补的细胞 tRNA 分子，在实验中，必须提供一种适当的引物。反转录酶最重要的用途是制备 cDNA 时从 RNA 合成 DNA，如第 5 章所述。反转录酶也可用于测序，只要能够得到一种合适的引物，就可以直接进行 RNA 测序，或 (较罕见) 直接进行 DNA 测序，因为反转录酶可以用 DNA 作模板，也可