



科爱传播
KE AI COMMUNICATIONS
生命科学

· 导读版 ·

Principles of Nucleic Acid Structure

核酸结构原理

Stephen Neidle



ELSEVIER
原版引进



科学出版社

www.sciencep.com

图字:01-2007-5843 号

This is an annotated version of
Principles of Nucleic Acid Structure by Stephen Neidle.

Copyright © 2008, Elsevier Inc.

ISBN 13:978-0-12-369507-9

All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

AUTHORIZED EDITION FOR SALE IN P. R. CHINA ONLY

本版本只限于在中华人民共和国境内销售

图书在版编目(CIP)数据

核酸结构原理=Principles of Nucleic Acid Structure:英文/(英)内德尔(Neidle, S.)主编. —影印本. —北京:科学出版社,2008

ISBN 978-7-03-021729-5

I. 核… II. 内… III. 核酸-分子结构-英文 IV. Q52

中国版本图书馆 CIP 数据核字(2008)第 056799 号

责任编辑:孙红梅 李小汀

责任印制:钱玉芬/封面设计:耕者设计工作室

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>

中国科学院印刷厂印刷

科学出版社发行 各地新华书店经销

*

2008年5月第 一 版 开本:787×1092 1/16

2008年5月第一次印刷 印张:19 1/2

印数:1—2 000 字数:462 000

定价:76.00 元

(如有印装质量问题,我社负责调换〈科印〉)

《核酸结构原理》——核酸结构生物学的优秀参考书

回顾历史,核酸结构研究曾经立下过丰功伟绩。1944年,Avery等人发表了著名的肺炎球菌转化实验。他们发现从一种类型的肺炎球菌,制备得到的DNA可以引入另一类型的肺炎球菌内,并使后者的遗传性状变成前者。他们的结论是:核酸(此处指DNA),而不是蛋白质,才是遗传物质。但是,大多数人仍然不能接受DNA携带遗传信息这一结论。因为蛋白质在生命中无处不在,并且种类繁多、功能复杂。他们不相信结构单一的核酸能控制如此繁复的蛋白质。直到1953年,基于Franklin和Wilkins对DNA进行的X光衍射分析的实验数据,Watson与Crick提出了DNA双螺旋结构模型,合理地解释了DNA的结构数据,以及DNA作为遗传信息携带者如何进行复制与转录。此后,DNA携带遗传信息的观点才被广泛接受。DNA双螺旋结构的发现确立了分子生物学的地位,使其与具有200年历史的生物化学并驾齐驱,这两门学科也成为半个多世纪以来生物学的先行学科。

在现代分子生物学研究中,核酸结构研究就像是足球场上的临门一脚。双螺旋结构模型提出后,人们的视线转移到蛋白质的生物合成。但没有料到花了十多年的时间,到1966年才基本解决了这一问题:rRNA与多种蛋白质组成核糖体,与mRNA结合后开始肽链的合成;遗传信息通过tRNA与mRNA自身的反密码子和密码子来识别,由不同的tRNA携带相应的氨基酸依次进入核糖体,并连接成多肽链。但当时对蛋白质合成过程仍有一步尚不清楚:没有找到催化肽键形成的转肽酶。科研人员一次次尝试纯化该酶的努力均告失败。后来进行的生物进化模拟实验证明,核酸有可能进化出转肽酶的活性。然而,就像今人用古老的技术模拟建造金字塔一样——它只是有可能,却不能证明历史确实如此。另一实验证明,脱去蛋白达95%以上的核糖体仍有合成蛋白质的活性。但质疑的声音仍在:会不会恰恰就是这无法去除的微量蛋白,催化了肽键的合成?可是三十多年来,对转肽酶的研究始终在禁区边缘徘徊……直到1999年,T. Steitz等人利用X衍射分析对核糖体50S亚基晶体结构的研究发现,在肽键形成位点的周围,除核酸的电子云外,看不到任何蛋白质的电子云。因此,the ribosome is a ribozyme(“核糖体是核酶”,即转肽酶就是rRNA)。这项科学成就仿佛一记“世界波”,当T. Steitz在第18届国际tRNA研讨会(英国剑桥,2000年4月7日~12日)上宣布这个消息的声音刚落,会场上一片欢呼声、掌声和德国人庆祝时特有的踏脚声。笔者亲临现场的感受是,科学家的疯狂毫不亚于足球场上的球迷。

在当今的分子生物学研究中,核酸结构研究是不可或缺的重要部分。2006年,中科院上海生科院生化与分子生物学研究所进行结构生物学研究的丁建平教授,与笔者所在的从事分子生物学研究的实验室合作,首先完成了人色氨酸tRNA合成酶-牛色氨酸tRNA复合物的晶体结构研究。美国科学院院士Schimmel等人(他们在上世纪从事分子生物学研究,本世纪已基本全部转为结构生物学研究)几乎同时完成了相同的工作。我们的实验室在比较了两个结构图的差异后,发现因两实验室操作的不同,造成合成酶色氨酸活化中心结构的细微差异。在此基础上通过分子生物学实验,证明了在该区域一个

氨基酸的改变,可以使该合成酶活化氨基酸的方式从原来的不依赖于 tRNA 转变成依赖于 tRNA。我们的工作深化了对合成酶的认识,有可能为其进化研究提供线索。该论文于 2007 年发表在国际著名学术杂志 *Nucleic Acids Res* 上。

核酸结构研究逐渐应用于制药研发等产业领域。随着计算机科学的发展,通过计算机模拟小分子与核酸、小分子与蛋白质(包括核酸结合蛋白)结合的方法,药物设计得到了飞速发展。

结构生物学渐渐成为生物科学的生长点。如果没有结构生物学数据的支持,单纯的生化与分子生物学实验成果想要在高水平科学杂志上发表,难度越来越高。美国有资格培养博士研究生的大学,几乎都有结构生物学实验室。目前,我国的分子生物学有了较大的发展,但与国际水平相比还有相当大的差距。在结构生物学方面的差距则更大,表现在能开展结构生物学研究的实验室很少,而进行核酸结构生物学研究的实验室则少之又少。国内没有专门介绍核酸结构的专著,有关核酸结构的介绍散在于生化教科书及相关专著中,无论是深度、广度,还是前沿性,都有不足。近年较为详尽的专著,只有笔者的《核糖核酸与核糖核酸组学》(2005,科学出版社),该书用一章的篇幅详细介绍了核酸的组成和结构元件等,如列举了组成核酸的一百多种天然核苷酸和在核酸中发现的一百多种碱基配对方式,以及三链核酸与四链核酸的形成原理等。但与将要介绍给读者的这本《核酸结构原理》相比,内容不够全面,也未列出任何晶体结构图。国内的学术刊物也少有结构生物学方面的论文发表,因为较高水平的论文多在国际杂志上发表了。因此,我国大生物领域的本科生、研究生必须从国际图书与杂志上获取结构生物学知识。相信 Neidle 教授的《核酸结构原理》一书,能为我国生物学界了解核酸结构原理和此领域最新的研究成果打开一扇窗户。在互联网上可检索到关于核酸结构的外文专著,但近几年来数量不多,更凸显出本书的重要性。

Stephen Neidle 毕业于伦敦的帝国学院化学系,获结晶学哲学博士学位。此后在剑桥的国王学院生物物理系从事博士后研究,1990 年任该研究所生物物理学教授(英国的教授是定编的,通常一个学科一个教授,有缺才补)。2002 年起任伦敦大学药学院的化学生物学教授,英国癌症研究与生物分子结构组主任。2004 年任皇家化学会化学生物学论坛主席,癌症医学中心主任。他曾获 2000 年度皇家化学会的生物与医学奖、2002 年度的学科交叉奖、法国化学治疗协会 2004 年度的 Paul Ehrlich 奖,以及医学化学的 Aventis 奖等。

Neidle 教授长期从事 X-光晶体学、药物-DNA 相互作用、推理抗癌药物设计、核酸结构与化学生物学研究。他是一位高产的科学家,每年发表研究论文 6~7 篇。如今年截至 4 月初已发表了 4 篇。他在 *Nature*, *Nature Rev Cancer*, *J American Chem. Society* 等国际顶尖杂志上发表了大量论文。同时,他又是一位高产的作家。从 1983 年起,他至少发表了以个人名义写作的 DNA 结构方面的专著 10 本。根据互联网检索,他的专著引用情况较好。Neidle 教授撰写的 *Principles of Nucleic Acids Structure*(核酸结构原理,2008 年),是在 *Nucleic Acid Structure and Recognition*(核酸结构及识别,2001 年)一书的基础上扩充而成。正如前言中所说,自那本书发表以后,在核酸结构方面有太多的进展,已经不能通过再版解决扩充内容的问题,所以只能“从头合成”。本书加入了核酸结构领域在 21 世纪最初 7 年中的新进展,也包括作者本人的研究成果。

从整体的角度来看,本书的特点如下:

(1)系统性。该书共分七章,分别介绍核酸结构的研究方法、核酸的组成,以及各种核酸的结构,使读者获得核酸结构研究的简明原理与方法,包括 X-光晶体学、核磁共振(NMR)、分子建模和数据库的讨论。

(2)前沿性。本书涵盖了 2000 年至 2007 年获得的最新研究成果,包括大量核酸结构图,以及这几年才开始出现或成为热点的内容,如核糖开关、小干扰 RNA、三链核酸、四链核酸等。

(3)新颖性。分子结构方面的图书都存在无法充分传递三维结构信息的缺点。而现在网络资源提供很多免费的优秀作图软件,从蛋白质数据库和其他网络站点可得到分子显示工具。因此作者在本书中给出了大量代表性结构的 PDB(蛋白质数据库)和 NDB(核酸数据库)码,使读者能快速在网上观看特殊的结构特征,也可下载结构文档,供以后在台式电脑或实验室平台上显示或分析。

(4)易读性。每章开篇都简述了与该章内容有关的背景知识、相关原理等,以帮助有一定基础的读者回忆起相关内容。每章的结尾还提供了大量参考文献,说明材料的出处。特别是分类列出相关的经典文献和最新文献,使得没有核酸化学或结构生物学背景的读者易于查找资料,在扩展阅读后更快地进入作者所叙述的世界。

(5)实用性。本书包括了重要并且具有普遍特征的结构,特别是与化学、生物学或药理学有关的结构图两百多幅,相关表格 60 余个。这些图表有很强的资料性,可供读者查阅和分析,从而很容易将所学的知识应用到化学、生物学及药理学等相关领域的学习、科研或教学工作中。

因此,本书是一本结构生物学领域的优秀参考书。读者群包括对结构生物学感兴趣的生物学各学科(特别是生物化学、分子生物学、化学生物学领域)、医学和药学领域(特别是化学药物设计领域)的在校本科生、研究生、博士后,各级教辅人员,以及上述领域的研究人员。

在当前浮躁的时代背景下,科研工作也常有追赶潮流的现象。一些人永远追逐即时的热点,对基础冷僻的知识则不愿花费功夫去掌握,这些方面常常会成为教学的盲点。根据笔者多年来为研究生授课及考核的经验,估计对一部分读者而言,本书最大的难点在于:对核酸结构式的生疏,使得全书两百余张结构图难于理解,读完全书会有留牍还珠的可能。唯有静下心来,认真阅读,多看参考,仔细揣摩,才有可能真正领悟。

为了帮助读者理解本书内容,将各章的要点或者切入点以引言的形式整理如下:

第一章 本章简要介绍核酸结构研究中两种主要的方法,侧重于它们的使用范围和局限。研究核酸结构的方法首推 X-光晶体结构分析,它提供了我们现在所知的、绝大多数的结构信息。此外,核磁共振谱、分子建模/模拟和化学/生化探针技术,对于提供结构、动力学信息和结构的柔性也非常重要。传统的光谱学方法可以提供重要且互补的巨观层面信息,而前面提到的那些方法可以在原子层面提供某些细节。近期发展的诸如表面细胞基质共振谱、单分子方法等,填补了巨观层面与原子层面之间的信息缺口。同时,核酸研究方法也得益于:1. 常规化学方法的进展,如毫克级 DNA、RNA 寡聚核苷酸的合成与纯化;2. RNA、DNA 结合蛋白,以及天然 RNA 分子的高效常规的克隆和表达系统方面的进展。

第二章 20 世纪早期的核酸降解研究表明,高分子量的核酸是由单独的酸性单元组

成。四种核苷酸(鸟嘌呤核苷酸、腺嘌呤核苷酸、胞嘧啶核苷酸和胸腺嘧啶核苷酸)已被分离出来,它们能被进一步剪切成磷酸根和四种核苷。核苷又由脱氧五碳糖和四种杂环碱基中的一种组成。碱基可以分为嘌呤碱基(腺嘌呤、鸟嘌呤)和嘧啶碱基(胸腺嘧啶、胞嘧啶)两类。核糖核酸中,胸腺嘧啶被尿嘧啶取代,尿嘧啶糖基的 2' 位上也带一羟基。核酸中原子的标准命名是由国际生化联盟批准的,见图 2.1 和 2.2。X-光晶体结构分析和测量已得到了精确的键长和键角参数的平均值。它们被编入 AMBER 和 CHARMM 力场的执行程序,广泛用于分子机制和动力学模拟,以及晶体和 NMR 结构分析的各种计算机软件包。极高分辨率的精确的晶体结构分析,也能直接给出分子中电子云密度分布的定量信息,从而给出了单个原子的电荷信息。这些数据最初由量子力学计算获得,现在也可通过实验获得。

单个的核苷通过核糖 3'、5' 位的磷酸根,相互连接成核酸。因此,核酸的重复单元是 3' 或 5' 核苷酸。核酸或寡核苷酸的序列用五种核苷酸的单字符 A、T、G、C、U 表示。两类核苷酸可以用 Y(嘧啶)和 R(嘌呤)表示。磷酸根常用小写的 p 表示。寡核苷酸链从 5' 开始排序。如 ApGpCpTpTpG 的 5' 端带有游离羟基的腺嘌呤核苷,和 3' 端带有游离羟基的鸟嘌呤核苷。字母 p 为居间的磷酸根,有时可以省略不写。涉及结构的文献中,常在一序列前加“d”,如 d(CGAT),以强调它是脱氧寡核苷酸,而不是核糖寡核苷酸。也常用前缀“r”表示核糖核苷、核糖核苷酸和寡核糖核苷酸。糖与碱基间的键为糖苷键。糖的立体化学非常重要。天然核酸中的糖苷键总是 β 型。从糖平面的方向看时,碱基在糖平面的上方。由此,碱基与 5' 取代物处于平面的同一面。当糖环氧原子 O4' 处于后排时,3' 位点的羟基处于环的下方,4' 位点的羟基处于上方。从化学的角度讲,有可能构建 α -核苷,由此可得到 α -核苷酸。这时,对应于糖环和其他基团,碱基处于下方。与天然 β -寡核苷酸相比,这种构型更能抵御核酸酶的攻击,因而具有超强的细胞内稳定性,并被用于 mRNA 的反义寡核苷酸。本章在此基础上介绍了 DNA 和 RNA 的组成,包括碱基配对、碱基和碱基对的柔性、糖折叠、糖苷键的构象、骨架的扭转角和柔性等内容。

第三章 X-光衍射谱分析可获得螺旋的基本参数,即重复单元长宽高数据。螺距(helical pitch, P)定义为一个完整螺旋单元在平行于螺旋轴方向的长度。当纤维垂直于 X-光光束时,会产生子午线反射,预示着在这个方向,存在结构的周期性。分析这种周期性反射,可得到有规律的重复距离 d(如果核苷酸单元垂直于螺旋轴,即螺旋的升距)。单元重复 n,是完整螺旋单元中核苷酸的数量,它等于 P/d 。如果核苷酸并不严格平行于螺旋轴,则螺旋升距为投射在螺旋轴上的连续两个核苷酸间的距离。螺旋的直径可从垂直于螺线管平面的投影数据中得到。

碱基对并不需要处于同一平面,事实上也鲜有处于同一平面的碱基对。部分原因是由于氢键的几何学需要并不严格。供体-H-受体间的角度可以从线性到 35° ,而不明显丧失氢键的能量。也有其他因素存在,如螺旋中某些碱基对-碱基对的非共价键相互作用中需要避免空间碰撞,使碱基对的两个碱基不在一个平面。这种背离同一平面的情况能够以多种方式发生,可归纳为两类:第一类是碱基对中一个碱基相对于另一个碱基的运动(如扭曲,“twist”和弯曲,“buckle”),碱基对相对于螺旋轴的移动;第二类是连续碱基对(碱基对台阶,“step”)中,两个碱基对的相对移动(如螺旋的扭曲,“twist”和摇摆,“roll”),它定义为结构中连续碱基对的相对关系。第二类方式不可避免地会伴随第一类

方式。本章着重介绍由此而形成的各种 DNA 结构。

第四章 碱基间的配对并不仅限于标准的沃森-克里克碱基对,还有各种非标准配对,如 G·G 等错配、相同碱基间反式-反式与反式-顺式的不同配对、碱基-烷基及碱基-糖基间的配对等。此外,自然界不仅存在双螺旋,还存在三链螺旋和四链结构等,而四链结构可以是单链 DNA 的折叠、双链 DNA 的折叠、四条单链间的多种方式形成。因此近年来发现,DNA 的结构非常复杂,种类极为丰富。本章主要介绍这方面的最新进展。

第五章 小分子、蛋白质和药物与 DNA 的识别,可以分为非特异性识别和特异性识别。最典型的非特异性识别是水与 DNA 间的识别,水可以与碱基上的氨基、酮基、杂环中的氮原子、脱氧核糖、磷酸根等形成氢键。这类识别不牵涉到对碱基对或核酸序列的识别。特异性识别则牵涉到对核酸序列的识别。根据四种核苷酸在序列中的随机分布原理,可以算出一定长度 DNA 识别序列出现的频率。因此在 32 亿个核苷酸长度的人类基因组中,只有 16~18 个核苷酸的序列才可能是独一无二的识别位点。同时,只有进入双螺旋沟的小分子等才能识别处于双螺旋内部的碱基对。通常假设 B-DNA 的小沟比大沟狭窄,因此认为小沟内 G·C、C·G、A·T 和 T·A 中的 3 个或 2 个氢键不大可能同时满足直接的氢键序列阅读。但是,沟内的结构水可显著增大沟的宽度,合适的配体可以稳定变宽的沟。其他序列性特征也可处于螺旋式扭曲和螺旋扭曲区以及长程影响区(如 A-片段区)。大沟序列的阅读中,胸腺嘧啶上的甲基起重要作用,它可以与亮氨酸、异亮氨酸和丙氨酸这些疏水氨基酸侧链形成稳定的范德华尔斯非共价键结合。当存在特殊侧链基团等情况时,可能发生立体碰撞造成 DNA-配体的失稳定性。因此,大量的胸腺嘧啶碱基可以驱动大沟中 G/C 区的识别。小沟中大量鸟嘌呤环外的 2-氨基是很多小沟配体的干扰物。然后结合更趋向于发生在 A/T 区。本章着重介绍了各种小分子与 DNA 结合的实例。

第六章 RNA 晶体与 NMR 结构研究在最近几年获得了快速发展。主要得益于长度为 15~20 聚 RNA 的毫克级合成与纯化技术的进展。更长的 RNA 可通过 T7RNA 聚合酶体外转录常规制备,尽管需要 5' 端有嘌呤序列以提高转录效率。传统的 RNA 结晶工作很困难,部分原因是 RNA 分子的自身特性。一系列新方法的产生,如矩阵方法、机器人操作、重原子或衍生物的导入、含衍生物的寡聚核苷酸的使用、特别是硒的参入等,已经可以成功做到使用少量 RNA 尝试大量的结晶条件,并较快得到大量 RNA 的结晶。

DNA 通常是以双链形式存在, RNA 则常以单链形式存在,虽然在 RNA 病毒中有可能有双链形式的。单链的 RNA 有很多种。mRNA 携带从基因序列传递来的遗传信息,进入由 rRNA 与多种蛋白质组成的核糖体,参与蛋白质的生物合成。tRNA 则参与 mRNA 的解码,对应于 20 种氨基酸,有 20 种 tRNA。另一类 RNA 是具酶活性的 RNA——核酶。它们能自行切割自身片段,可能是在 pre-DNA world,即前 DNA 世界起主要功能的分子。近来还发现了一类复杂序列的 RNA,它们卷曲形成分子开关(核糖开关, riboswitch),可以结合某些小分子化合物。还有一类 20~25 个核苷酸的双链 RNA,小干扰 RNA(siRNA)。它们通过 RNAi 机制能抑制真核基因的表达。本章着重介绍上述各类 RNA 结构。

第七章 DNA 携带的遗传信息在参与复制、转录和翻译的过程中,需要与大量的非特异性及特异性蛋白相互作用。DNA 的损伤应答也需要与大量蛋白质发生互作。因此,有关 DNA-蛋白相互作用的研究得到快速发展并不意外,这方面的研究也逐渐渗透到生物学的各个分支领域。大量的结构知识支撑着生物学与生物物理学的发展,这些知识主要来自晶体结构研究及 NMR 结构研究。有大量真核生物和原核生物的各种 DNA-蛋白的结构数据可被利用,数据量还在不断增加。1997 年 PDB 和 NDB 数据库有 241 个结构,而到 2007 年 6 月已增加到 1789 个,当然相对于基因组编码的数量,这个数字还很小。现在,主要的努力都集中在大规模高通量的结构研究(结构基因组学)上。但是我们对 DNA-蛋白识别的理解不仅来自于单个蛋白的复合物,更多的是来自与功能有关的多蛋白复合物,如与翻译有关的核糖体和与转录有关的聚合酶复合物。按照功能和结构可以将 DNA-结合蛋白分为几个大类。如按功能可分为:调控蛋白、DNA 切割蛋白(核酸酶)、修复蛋白、复制前解决各种拓扑学问题的蛋白、结构蛋白、加工蛋白。相对于 DNA-结合蛋白的功能多样性,可被识别的 DNA 结构较少。但是 DNA 结合蛋白本身的结构变化很大。大沟通常是直接识别信息的场所,而小沟则是调控蛋白和结构蛋白的靶区,特别是能使 DNA 变形的蛋白,以致于小沟常被大大扩展。

我们通常假设 B-DNA 中都是沃森-克里克碱基对,因此蛋白质必须识别沃森-克里克碱基对。但在已鉴定的少量 DNA-蛋白结构中,已观察到 Hoogsteen 碱基对。在 DNA-结合蛋白中发现的识别结构有: α -螺旋-转折- α -螺旋(HTH 结构)、锌指和环、拉链结构、 β -折叠和 β -发夹。本章重点介绍各种最新的蛋白质-DNA 识别实例。

金由辛

中国科学院上海生命科学研究院
生物化学与分子生物学研究所
yxjin@sibs.ac.cn

前 言

自前一本书于 2002 年出版以来(译者注:指作者于 2001 年写的 *Nucleic Acid Structure and Recognition*,《核酸结构及识别》),时光消逝又几年,那本书已进入核酸研究方面的重要参考书之列。2003 年,我们庆祝了 DNA 双螺旋结构发现 50 周年和人类基因组序列的发表。由此,您可能认为核酸结构本身已成为历史的一部分,已经没有什么是不知道的。而事实并非如此。在本世纪的前七年中,很多与 RNA、DNA 结构有关的新发现意义深远。这些重要的进展,使原书的很多部分需要扩充,并重新写作其他的部分。这已远非再版所能及。

本书的目的是提供有关核酸结构的基本特征、原理和很多核酸结构的实例。希望为读者继续学习结构生物学和核酸化学提供坚实的基础。它的读者群是希望在研究中运用相关知识的研究生,以及对核酸结构感兴趣的高年级本科生。本书不打算全面阐述所有包含核酸的结构,而是涉及大多数读者关心的问题,并聚焦于这样一些结构:能代表重要、普遍特征的结构,特别是与化学、生物学或药理学问题有关的结构。如因上述主观判断使得本书未能涉及某些读者感兴趣的结构,我在此先表歉意。

本书着重于 X-光晶体学测得的结构,因为这种方法在按分子大小测定结构的领域中占主导地位,并依然为我们提供着绝大部分的高精度结构。书中介绍晶体学和其他技术的目的,是帮助非专业人士在理解的基础上阅读原始文献,更重要的是能够判断实验与理论结构研究的范围和特点。我扩展了参考文献和读书目录的部分,也包括很多相关网站的信息,以提供对各个时期的文献合理而全面的指南。

分子结构方面的专著都存在不能充分传递三维结构信息的缺点。2002 年的那本书提供了专用网站的站点,读者可以通过多种显示方式对结构进行交互检查。现在,网络提供很多免费的优秀作图软件,从蛋白质数据库和其他网络站点可以得到分子显示工具。因此提供专用站点已不再是必需的,甚至有点不合时宜。我已经为很多具有代表性的结构提供了 PDB(蛋白质数据库)和 NDB(核酸数据库)代码的表格,以帮助读者快速观看特殊的结构特征或下载结构文档,可供读者在自己的台式电脑或实验室平台上显示或分析。我也提供了自己偏爱的分子图象软件,它们具有支持核酸结构的运行特征。

感谢我的妻子 Andrea 和我的孩子 Dan, Ben, Hannah,在我写作此书的过程中及其他工作中所给予的一贯支持和鼓励。感谢我的同事、合作者和学生对此书的贡献、洞察力和讨论。以及 Elsevier 出版集团的编辑 Kirsten Funk,感谢她的辛勤工作、耐心和给予我的支持。

Stephen Neidle
2007 年 6 月于伦敦

目 录

前言	vii
目录	ix
1. 核酸结构的研究方法	1
1.1 引言	1
1.2 用于结构分析的 X-光衍射法	2
1.2.1 概况	2
1.2.2 纤维衍射法	5
1.2.3 单晶法	7
1.3 研究核酸结构与动力学的核磁共振(NMR)法	10
1.4 核酸的分子建模与模拟	11
1.5 结构与动力学的化学、酶学和生物物理探针	14
1.6 结构数据源	15
1.7 核酸分子结构的可视化	15
1.7.1 本书中的结构	16
2. DNA 和 RNA 的组成	20
2.1 引言	20
2.2 碱基配对	23
2.3 碱基和碱基对的柔性	24
2.4 糖折叠	28
2.5 糖苷键的构象	32
2.6 骨架的扭转角和相关柔性	33
3. 纤维和晶体中观察到的 DNA 结构	38
3.1 结构的基本原理	38
3.1.1 螺旋的参数	38
3.1.2 碱基对的形态学特征	38
3.2 纤维衍射研究所得到的多聚核苷酸结构	39
3.2.1 经典的 DNA 结构	39
3.2.2 纤维中 DNA 的多态性	43
3.3 晶体结构分析所看到的 B-DNA 寡聚核苷酸结构	47
3.3.1 Dickerson-Drew 十二聚	47
3.3.2 Dickerson-Drew 十二聚的其他研究	49
3.3.3 其他 B-DNA 寡聚核苷酸的结构	51

3.3.4	B-DNA 依赖于序列的特征:事例与预测	56
3.4	A-DNA 寡聚核苷酸的晶体结构	60
3.4.1	A-型八核苷酸	60
3.4.2	溶液中存在 A 型寡聚核苷酸吗? 晶体-堆积效应	61
3.4.3	晶体中 A $\leftrightarrow$ B 的转变	63
3.5	Z-DNA—左手 DNA	64
3.5.1	Z-DNA 六核苷酸晶体结构	64
3.5.2	大体的结构特征	65
3.5.3	Z-DNA 螺旋	66
3.5.4	其他 Z-DNA 结构	67
3.5.5	Z-DNA 的生物学问题	67
3.6	弯曲的 DNA	69
3.6.1	溶液中的 DNA 周期	69
3.6.2	A-片段及其弯曲	70
3.6.3	显示弯曲的结构	71
3.6.4	PolydA · dT 结构	73
3.7	小结	74
4.	DNA 的非标准结构和高级结构:DNA-DNA 识别	81
4.1	DNA 中的错配	81
4.1.1	一般特征	81
4.1.2	嘌呤:嘌呤错配	82
4.1.3	烷基化的错配	85
4.2	DNA 三链螺旋	88
4.2.1	导言	88
4.2.2	结构研究	90
4.2.3	反平行三链螺旋与非标准碱基配对	95
4.2.4	三链的应用	100
4.3	鸟嘌呤四链结构	101
4.3.1	导言	101
4.3.2	DNA 四链结构的总体特征	103
4.3.3	简单的四链结构	107
4.3.4	复杂的四链结构	108
4.3.5	i-模体	113
4.4	DNA 连接	114
4.4.1	Holliday 连接结构	114
4.4.2	DNA 酶的结构	118
4.5	非天然 DNA 的结构	120

5. 小分子-DNA 的识别原理	132
5.1 引言	132
5.2 DNA-水的相互作用	136
5.2.1 沟中水合作用的细节	140
5.3 DNA-药物、DNA-小分子识别的一般特征	143
5.4 插入结合	144
5.4.1 简单的插入物	146
5.4.2 复杂的插入物	147
5.4.3 大沟中的插入	151
5.4.4 双插入	158
5.5 DNA 高级结构中的嵌入型结合	163
5.5.1 三链 DNA-配体相互作用	163
5.5.2 配体与四链 DNA 的结合	164
5.5.3 配体与连接 DNA 的结合	166
5.6 与沟结合的分子	169
5.6.1 简单的沟结合分子	169
5.6.2 纺锤菌素和偏端霉素	178
5.6.3 序列专一性的聚酰胺	182
5.7 与 DNA 共价结合的小分子	187
5.7.1 铂药物	188
5.7.2 同时为序列特异性的和共价结合的小分子	191
 6. RNA 结构及其多样性	 204
6.1 引言	204
6.2 RNA 结构的基本原理	206
6.2.1 RNA 的螺旋构象	206
6.2.2 RNA 的错配与突起结构	210
6.3 tRNA 结构	217
6.4 核酶	221
6.4.1 锤头型核酶	223
6.4.2 复杂型核酶	224
6.5 核糖开关	227
6.6 核糖体,核酶机器	229
6.6.1 30S 亚基的结构	232
6.6.2 50S 亚基的结构	234
6.6.3 完整的核糖体结构	234
6.7 RNA-药物复合物	235
6.8 RNA 结构	241

7. 蛋白质-DNA 识别的原理	249
7.1 导言	249
7.2 蛋白质-DNA 的直接接触	252
7.3 与大沟的相互作用—— α 螺旋作为识别元件	257
7.4 锌指的识别模式	259
7.5 其他大沟识别结构	263
7.6 与小沟的识别	264
7.6.1 B-DNA 的识别	264
7.6.2 由 TBP 打开小沟	267
7.6.3 诱导 DNA 弯曲的其他蛋白质	268
7.7 DNA 弯曲与蛋白质识别	272
7.8 蛋白质-DNA-小分子识别	275
索引	283

(金由辛 译)

Contents

1. Methods for Studying Nucleic Acid Structure	1
1.1 Introduction	1
1.2 X-ray Diffraction Methods for Structural Analysis	2
1.2.1 Overview	2
1.2.2 Fiber Diffraction Methods	5
1.2.3 Single-Crystal Methods	7
1.3 NMR Methods for Studying Nucleic Acid Structure and Dynamics	10
1.4 Molecular Modelling and Simulation of Nucleic Acids	11
1.5 Chemical, Enzymatic, and Biophysical Probes of Structure and Dynamics	14
1.6 Sources of Structural Data	15
1.7 Visualization of Nucleic Acid Molecular Structures	15
1.7.1 The Structures in This Book	16
 2. The Building-Blocks of DNA and RNA	 20
2.1 Introduction	20
2.2 Base Pairing	23
2.3 Base and Base Pair Flexibility	24
2.4 Sugar Puckers	28
2.5 Conformations About the Glycosidic Bond	32
2.6 The Backbone Torsion Angles and Correlated Flexibility	33
 3. DNA Structure as Observed in Fibers and Crystals	 38
3.1 Structural Fundamentals	38
3.1.1 Helical Parameters	38
3.1.2 Base-Pair Morphological Features	38
3.2 Polynucleotide Structures from Fiber Diffraction Studies	39

3.2.1 Classic DNA Structures	39
3.2.2 DNA Polymorphism in Fibers	43
3.3 B-DNA Oligonucleotide Structure as Seen in Crystallographic Analyses	47
3.3.1 The Dickerson–Drew Dodecamer	47
3.3.2 Other Studies of the Dickerson–Drew Dodecamer	49
3.3.3 Other B-DNA Oligonucleotide Structures	51
3.3.4 Sequence-Dependent Features of B-DNA: Their Occurrence and Their Prediction	56
3.4 A-DNA Oligonucleotide Crystal Structures	60
3.4.1 A-Form Octanucleotides	60
3.4.2 Do A-Form Oligonucleotides Occur in Solution? Crystal-Packing Effects	61
3.4.3 The A $\leftrightarrow$ B Transition in Crystals	63
3.5 Z-DNA – Left-Handed DNA	64
3.5.1 The Z-DNA Hexanucleotide Crystal Structure	64
3.5.2 Overall Structural Features	65
3.5.3 The Z-DNA Helix	66
3.5.4 Other Z-DNA Structures	67
3.5.5 Biological Aspects of Z-DNA	67
3.6 Bent DNA	69
3.6.1 DNA Periodicity in Solution	69
3.6.2 A-Tracts and Bending	70
3.6.3 Structures Showing Bending	71
3.6.4 The Structure of Poly dA•dT	73
3.7 Concluding Remarks	74
4. Nonstandard and Higher-Order DNA Structures: DNA–DNA Recognition	81
4.1 Mismatches in DNA	81
4.1.1 General Features	81
4.1.2 Purine:Purine Mismatches	82
4.1.3 Alkylation Mismatches	85
4.2 DNA Triple Helices	88
4.2.1 Introduction	88
4.2.2 Structural Studies	90
4.2.3 Antiparallel Triplexes and Nonstandard Base-pairings	95
4.2.4 Triplex Applications	100
4.3 Guanine Quadruplexes	101
4.3.1 Introduction	101
4.3.2 Overall Structural Features of Quadruplex DNA	103
4.3.3 Examples of Simple Quadruplex Structures	107

4.3.4 Some Complex Quadruplex Structures	108
4.3.5 The i-Motif	113
4.4 DNA Junctions	114
4.4.1 Holliday Junction Structures	114
4.4.2 DNA Enzyme Structures	118
4.5 Unnatural DNA Structures	120
5. Principles of Small Molecule-DNA Recognition	132
5.1 Introduction	132
5.2 DNA-Water Interactions	136
5.2.1 Hydration in the Grooves in Detail	140
5.3 General Features of DNA-Drug and Small-Molecule Recognition	143
5.4 Intercalative Binding	144
5.4.1 Simple Intercalators	146
5.4.2 Complex Intercalators	147
5.4.3 Major-Groove Intercalation	151
5.4.4 Bis-Intercalators	158
5.5 Intercalative-Type Binding to Higher-Order DNAs	163
5.5.1 Triplex DNA-Ligand Interactions	163
5.5.2 Ligand Binding to Quadruplex DNAs	164
5.5.3 Ligand Binding to Junction DNAs	166
5.6 Groove-Binding Molecules	169
5.6.1 Simple Groove Binding Molecules	169
5.6.2 Netropsin and Distamycin	178
5.6.3 Sequence-Specific Polyamides	182
5.7 Small Molecule Covalent Bonding to DNA	187
5.7.1 The Platinum Drugs	188
5.7.2 Covalent-Binding Combined with Sequence-Specific Recognition	191
6. RNA Structures and Their Diversity	204
6.1 Introduction	204
6.2 Fundamentals of RNA Structure	206
6.2.1 Helical RNA Conformations	206
6.2.2 Mismatched and Bulged RNA Structures	210
6.3 Transfer RNA Structures	217
6.4 Ribozymes	221
6.4.1 The Hammerhead Ribozyme	223
6.4.2 Complex Ribozymes	224
6.5 Riboswitches	227

6.6 The Ribosome, a Ribozyme Machine	229
6.6.1 The Structure of the 30S Subunit	232
6.6.2 The Structure of the 50S subunit	234
6.6.3 Complete Ribosome Structures	234
6.7 RNA-Drug Complexes	235
6.8 RNA Motifs	241
7. Principles of Protein-DNA Recognition	249
7.1 Introduction	249
7.2 Direct Protein-DNA Contacts	252
7.3 Major-Groove Interactions – the α -Helix as the Recognition Element	257
7.4 Zinc-Finger Recognition Modes	259
7.5 Other Major Groove Recognition Motifs	263
7.6 Minor-Groove Recognition	264
7.6.1 Recognition of B-DNA	264
7.6.2 The Opening-up of the Minor Groove by TBP	267
7.6.3 Other Proteins that Induce Bending of DNA	268
7.7 DNA-Bending and Protein Recognition	272
7.8 Protein-DNA-Small Molecule Recognition	275
Index	283