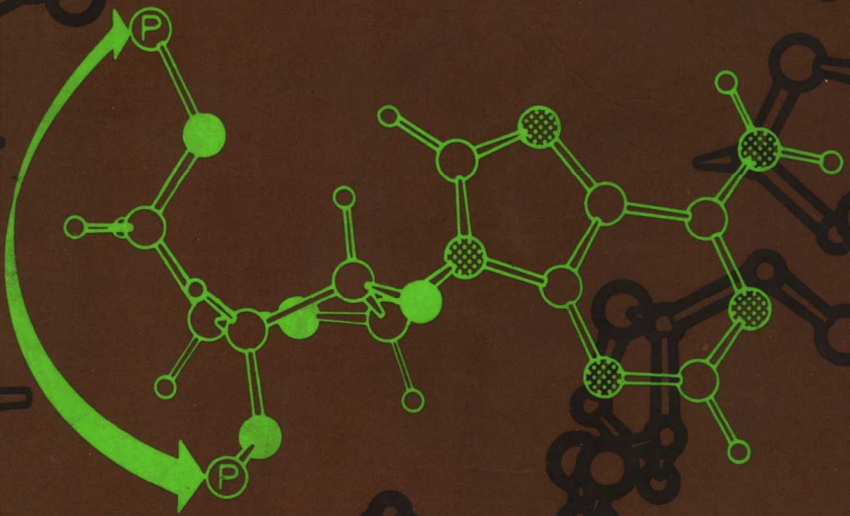
Wolfram Saenger

Principles of Nucleic Acid Structure



Springer-Verlag
New York Berlin Heidelberg Tokyo

Wolfram Saenger

Principles of Nucleic Acid Structure

With 227 Figures

Springer-Verlag
New York Berlin Heidelberg Tokyo

Wolfram Saenger
Institut für Kristallographie
Freie Universität
1000 Berlin 33
Federal Republic of Germany

Series Editor:

Charles R. Cantor
Department of Human Genetics and Development
College of Physicians & Surgeons of Columbia University
New York, New York 10032 U.S.A.

Library of Congress Cataloging in Publication Data
Saenger, Wolfram.
Principles of nucleic acid structure.
(Springer advanced texts in chemistry)
1. Nucleic acids. I. Title. II. Series.
QD433.S24 1983 547.7'9 82-19445

Typeset by Progressive Typographers, Emigsville, Pennsylvania.
Printed and bound by R.R. Donnelley & Sons, Crawfordsville, Indiana.
Printed in the United States of America.

9 8 7 6 5 4 3 2 1

ISBN 0-387-90762-9 Springer-Verlag New York Heidelberg Berlin (Hardcover Edition)
ISBN 3-540-90762-9 Springer-Verlag Berlin Heidelberg New York (Hardcover Edition)
ISBN 0-387-90761-0 Springer-Verlag New York Heidelberg Berlin (Softcover Edition)
ISBN 3-540-90761-0 Springer-Verlag Berlin Heidelberg New York (Softcover Edition)

Series Preface

New textbooks at all levels of chemistry appear with great regularity. Some fields like basic biochemistry, organic reaction mechanisms, and chemical thermodynamics are well represented by many excellent texts, and new or revised editions are published sufficiently often to keep up with progress in research. However, some areas of chemistry, especially many of those taught at the graduate level, suffer from a real lack of up-to-date textbooks. The most serious needs occur in fields that are rapidly changing. Textbooks in these subjects usually have to be written by scientists actually involved in the research which is advancing the field. It is not often easy to persuade such individuals to set time aside to help spread the knowledge they have accumulated. Our goal, in this series, is to pinpoint areas of chemistry where recent progress has outpaced what is covered in any available textbooks, and then seek out and persuade experts in these fields to produce relatively concise but instructive introductions to their fields. These should serve the needs of one semester or one quarter graduate courses in chemistry and biochemistry. In some cases the availability of texts in active research areas should help stimulate the creation of new courses.

New York

CHARLES R. CANTOR

Preface

This monograph is based on a review on polynucleotide structures written for a book series in 1976. The respective volume was, however, never published and the manuscript was kept in the drawer until Springer-Verlag decided to accept it as part of the Advanced Texts in Chemistry. When I looked at my 1976 manuscript again in 1980, I found that many views had changed. Left-handed DNA, above all, had been introduced and changed the notion of the occurrence of exclusively right-handed double-helices. As a matter of fact, the whole story had to be rewritten.

But even then, during the past two years many new discoveries were made and it often appeared to me that I was trying to catch a running train. Research on DNA structure is now booming, caused in part by the availability of synthetic oligonucleotides of defined sequence. They can be crystallized either per se or as complexes with certain metal ions, drugs, and proteins binding to them. And with the proteins again a new field has opened. We can now look at the three-dimensional structures of quite a number of DNA- and RNA-binding proteins. Work is in progress with more complicated systems involving the respective complexes, made possible by recent advances in crystallographic techniques.

I cannot close this short preface without thanking those who have contributed enthusiastically to this book. There are, above all, Meinhard Heidrich and Manfred Steifa who operated the Cambridge Data File to retrieve atomic coordinates. These were used in a combination of programs SCHAKAL (Ref. 1376) and ORTEP (Ref. 1377) to do most of the artistic plots of molecules displayed in this book. Drawings which could not be prepared on a computer came from the hand of Ludwig Kolb who did a magnificent and superb job on the rough drawings I gave him.

If writing a book, one has to have available a good library and people

associated with it. The library at the Max-Planck-Institut für Experimentelle Medizin is one of the best I have encountered and Mrs. M. Henschel, B. Küttner and Mr. W. Gellert helped to make available the more exotic journals. Finally, writing and rewriting many versions of the manuscript was done patiently by several secretaries, Chantal Celli, Jenny Glissmeyer and Petra Große contributing predominantly.

Last, but certainly not least, I would like to thank my scientific teacher Professor Friedrich Cramer for his guidance and support during the past 18 years.

Göttingen, August 1983 WOLFRAM SAENGER

Acknowledgments

Acknowledgment is gratefully made for permission to reproduce material as follows:

Figures 6-3, 6-15, 6-16, 6-17, 6-18, 11-4, 11-5, 11-6, 12-1, 12-3, 12-4, 12-8, 13-14, 15-2, 15-3, 16-3, 17-6, 17-7, and 18-11 are reprinted by permission from the *Journal of Molecular Biology,* copyright Academic Press Inc. (London) Ltd.

Table 3-2 is reprinted by permission from the *Journal of Theoretical Biology,* copyright Academic Press Inc. (London) Ltd.

Reprinted with permission from the *Journal of the American Chemical Society* are Figure 6-10, copyright 1967 American Chemical Society; Figures 2-8 and 15-10, copyright 1972 American Chemical Society; Figures 3-8 and 7-6, copyright 1977 American Chemical Society; Figures 4-6 and 13-13, copyright 1978 American Chemical Society.

Reprinted by permission from *Nature* are Figure 18-7, copyright 1955 Macmillan Journals Ltd; Figure 10-5, copyright 1971 Macmillan Journals Ltd; Table 6-10, copyright 1973 Macmillan Journals Ltd; Figure 4-24, copyright 1977 Macmillan Journals Ltd; Figure 4-29a and Table 4-4, copyright 1978 Macmillan Journals Ltd; Figure 17-10, copyright 1979 Macmillan Journals Ltd; Figures 11-3, 16-7, and 16-8, copyright 1980 Macmillan Journals Ltd; Figures 18-17 and 18-18, copyright 1981 Macmillan Journals Ltd; Figures 8-5 and 17-12, copyright 1982 Macmillan Journals Ltd.

Reprinted by permission from *Science* is Figure 1-3, copyright 1969 by the American Association for the Advancement of Science.

Reprinted by permission of the authors from *Proceedings of the National Academy of Science, U.S.A.* are Figures 3-5, 6-5, 6-6, 14-2, 14-3, 14-5, 16-6, 18-3, 18-5, 18-12, 19-2, and 19-7.

Figures 4-27, 6-14, 6-20, and 16-5 are reprinted by permission from *Biopolymers,* copyright John Wiley and Sons. Figure 15-1 is reprinted by permission from *Advances in Enzymology,* copyright John Wiley and Sons.

Figures 19-4, 19-6, 19-8 are from "The Nucleosome" by Roger D. Kornberg and Aaron Klug, copyright 1981 by Scientific American, Inc. All rights reserved. Box 19-3 is from "Supercoiled DNA" by William R. Bauer, F.H.C. Crick, and James H. White, copyright 1980 by Scientific American, Inc. All rights reserved.

Figure 19-1 is reprinted by permission from *Cell,* copyright M.I.T. Press.

Contents

Chapter 4

Structures and Conformational Properties of Bases, Furanose Sugars, and Phosphate Groups 51

Chapter 5

Physical Properties of Nucleotides: Charge Densities, p*K* Values, Spectra, and Tautomerism 105

Chapter 6

Forces Stabilizing Associations Between Bases: Hydrogen Bonding and Base Stacking 116

Chapter 1

Why Study Nucleotide and Nucleic Acid Structure?

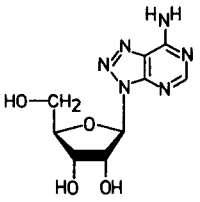
Before embarking on a description of nucleotide and nucleic acid structures, let us examine the biological importance of this class of molecules and find out why their structural principles should be known at the atomic level.

Nucleotides have many functions in living organisms. The hereditary material, deoxyribonucleic acid (DNA), is a linear polymer built up of monomeric units, the nucleotides. Even if these units were not constituents of DNA, they would nevertheless be among the most important molecules in biology.

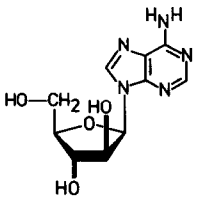
A nucleotide consists of three molecular fragments: sugar, heterocycle, and phosphate. The sugar, ribose or deoxyribose, is in a cyclic, furanoside form and is connected by a β-glycosyl linkage with one of four heterocyclic bases to produce the four normal nucleosides: adenosine, guanosine, cytidine, and thymidine (uridine in ribonucleic acid, RNA). If the 3′- or 5′-hydroxyl group of sugar is phosphorylated, we have a nucleotide. This unit, the nucleotide, is not only the building block of the polynucleotides DNA and RNA but it also exhibits independent functions.

For example, with adenosine derivatives displayed in Figure 1-1, we can show that, depending on chemical modifications, adenosine adapts to several, dramatically different biochemical roles in life. As di- and triphosphates, adenosine acts as an energy pool for many enzymatic processes and for muscle work. The importance of adenosine triphosphate is demonstrated by its turnover rate in humans: about one body weight per day per person. The 3′, 5′-cyclic phosphate of adenosine is the "second hormonal messenger," controlling and mediating the activities of peptide hormones. In the form of puromycin, adenosine is a potent inhibitor of protein biosynthesis and as arabino- or 8-azaderivatives, adenosines display antibiotic activities. Adenosine diphosphate, equipped at the terminal phosphate with certain biological molecules, is a constituent of both coenzymes A and NAD^+ and is essential for the proper functioning of enzymes which require these cofactors.

To understand the biological function of a nucleotide, we must know its structural features. With the adenosine derivatives mentioned above, we know the *chemical structure* and we know the function. But why and how do they function? Why, for instance, is 8-azaadenosine an antibiotic even though, from a chemical point of view, the isoelectronic substitution of

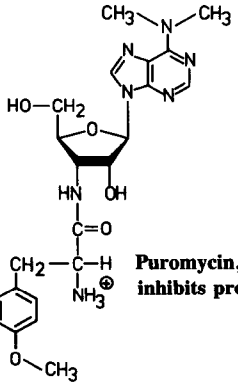


8-Azaadenosine

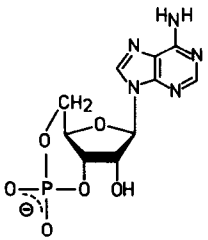


1-β-D-Arabinofuranosyl-adenine (araA)

Two potent antibiotics

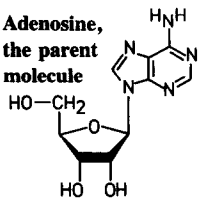


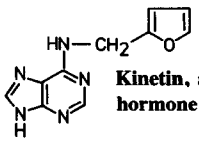
Puromycin, an antibiotic that inhibits protein biosynthesis



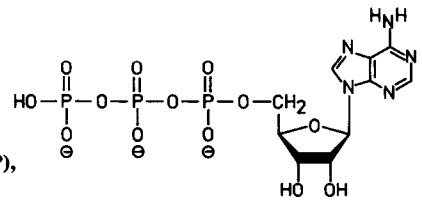
Cyclic 3′,5′-adenosine monophosphate (cAMP), the "second hormonal messenger"

Adenosine, the parent molecule

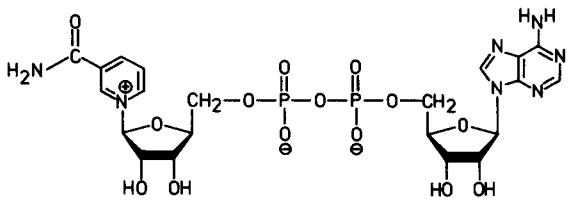




Kinetin, a plant growth hormone



Adenosine triphosphate (ATP), the energy pool



Nicotinamide-adenine-dinucleotide (NAD^+), a coenzyme

Figure 1-1. The properties of a biological molecule can change dramatically with only minor chemical modifications. This scheme displays adenosine (color) and some of its derivatives which exhibit very different biochemical behavior.

8-CH by an 8-aza group is expected to have only a modest influence on its chemistry? In this case it is not the electronic but rather the *three-dimensional structure* or conformation of the nucleoside which is changed and which explains 8-azaadenosine's particular properties.

Because nucleosides and nucleotides interact with proteins in all their metabolic or control operations, specific, mutual recognition of the two (or more) reactants is required. This presupposes that the partners involved have well-defined three-dimensional structures which, if we desire