PHYSICO-CHEMICAL PROPERTIES OF NUCLEIC ACIDS

Volume 2

edited by Jules Duchesne

Academic Press London and New York

A Subsidiary of Harcourt Brace Jovanovich, Publishers

Physico-chemical Properties of Nucleic Acids

edited by J. DUCHESNE
Department of Molecular
and Atomic Physics
University of Liège
Belgium

VOLUME 2: *Structural studies on nucleic acids and other biopolymers*

ACADEMIC PRESS INC. (LONDON) LTD.
24/28 Oval Road
London NW1

United States Edition published by
ACADEMIC PRESS INC.
111 Fifth Avenue
New York, New York 10003

Library of Congress Catalog Card Number: 72–84360
ISBN: 0 12 222902 9

PRINTED IN GREAT BRITAIN BY
Adlard & Son Ltd., Bartholomew Press, Dorking

Contributors to Volume 2

J. BRAHMS, *Institut de Biologie Moléculaire de la Faculté des Sciences, Paris, France.*
C. A. BUSH, *Department of Chemistry, Illinois Institute of Technology, Chicago, Illinois, U.S.A.*
S. I. CHAN, *A. A. Noyes Laboratory of Chemical Physics, California Institute of Technology, Pasadena, California, U.S.A.*
J. E. CRAWFORD, *A. A. Noyes Laboratory of Chemical Physics, California Institute of Technology, Pasadena, California, U.S.A.*
W. FIERS, *Laboratory of Molecular Biology, University of Ghent, Belgium.*
I. HARADA, *Department of Chemistry, Faculty of Science, University of Tokyo, Japan.*
K. A. HARTMAN, *Department of Microbiology and Biophysics, University of Rhode Island, Kingston, Rhode Island, U.S.A.*
R. C. LORD, *Department of Chemistry and Spectroscopy Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.*
K. E. REINERT, *Akademie der Wissenschaften der D.D.R. Zentralinstitut für Mikrobiologie und experimentelle Therapie, Abteilung Biophysikochemiie, DDR–69, Jena*
D. RIESNER, *Gesellschaft für Molekularbiologische Forschung m.b.H., 3301 Stoeckheim/Braunschweig, Germany.*
R. ROMER, *Gesellschaft für Molekularbiologische Forschung m.b.H., 3301 Stoeckheim/Braunschweig, Germany.*
M. P. SCHWEIZER, *ICN Nucleic Acid Research Institute, Irvine, California, U.S.A.*
S. TAKAHASHI, *Faculty of Pharmaceutical Sciences, University of Tokyo, Japan.*
G. J. THOMAS, jr., *Department of Chemistry, South-Eastern Massachusetts University, North Dartmouth, Massachusetts, U.S.A.*
M. TSUBOI, *Faculty of Pharmaceutical Sciences, University of Tokyo, Japan.*

Preface

Progress in the application of physics and chemistry to biology has been particularly important in a number of areas, and this book brings together recent work in these fields.

Nearly all the topics have been approached in the spirit of molecular biology and biophysics, where recent advances are far from having exhausted their impact.

Though the book has been subdivided into three parts for practical reasons, the unity of the whole work has been largely preserved and each volume is also an entity as it stands.

The first volume is mainly concerned with the intrinsic properties of nucleic acids, considered as macromolecules, and their components, and this study is completed by the analysis of different types of bindings or interaction mechanisms, including photodynamic and radiation effects, as well as fluorescence. In the second volume, the emphasis is put on structural studies and especially on conformational changes, using spectroscopic techniques as well as methods of thermodynamics and hydrodynamics. The stage of specific biological functions is attained in the last volume, with some considerations on repair mechanisms in relation to the general problem of evolution.

"The Physico-chemical Properties of Nucleic Acids" is intended to provide thought provoking material for research scientists, whether they are biologists, chemists or physicists. At the same time, it should be a source of information and reference for graduate students in these fields.

It is particularly hoped that these papers may help to stimulate the search for a better understanding of the correlation between structure and function. This understanding can only be founded on a detailed knowledge of the molecular properties of the basic substances.

Lastly, it is a great pleasure for the Editor to acknowledge the excellent cooperation of the publisher, who has been helpful in every way.

November 1972 JULES DUCHESNE

Foreword

The chemist, if he encounters a new substance, is accustomed to look first for its macroscopic properties in order to correlate them with underlying structural features. Such properties instantaneously suggest to him some picture of the molecule, and often it is only the material quality which stimulated his interest in particular substances.

If we think of nucleic acids, such a correlation does not usually come to our mind. In fact, most of us working on one or the other aspect of this fascinating molecule, have not even seen any pure crystalline material, not to speak of particular samples such as isolated gene material or single uniform *t*RNA batches, which actually are the objects of the most exciting studies in molecular biology.

Speaking of nucleic acids we usually do not associate with them any characteristic material property: we rather think of some abstract quality: information, instruction, translation, etc. We see before us sequences of letters, such as

$$\ldots \text{pApUpGpCpGpUpApUpApGpCpApUp} \ldots$$

and we think of a message. Of course, this is an exaggeration. Most of us—encountering any analysed sequence—would immediately start to play around with it and fold it up, e.g. as

$$\begin{array}{r} \ldots \text{pApUpGpC}^{\text{pG}}\text{pU} \\ \text{pA} \\ \ldots \text{pUpApCpG}_{\text{pA}}\text{pU} \end{array}$$

This shows that we actually associate with each letter also a particular physical interaction unique to this species of macromolecules.

It is this interaction, this exclusive way of complementary pairing of bases which is behind the abstract property of "code reading" or "information transfer", and it was the realisation of this quality which led Watson and Crick to their epochal discovery (quantitatively manifested in Wilkin's X-ray diffraction data).

How far our interpretations are guided by abstract reasoning became especially apparent in the deciphering of the genetic code. With 20 amino acids (plus some punctuation symbols) to be coded by the four bases (i.e. A, U (or T), G and C) it was "obvious" that the code had to be a triplet

code—as indeed has been confirmed by the work of Nirenberg, Matthaei, Khorana and Ochoa. However, this was "obvious" only by logical arguments: a doublet code could only provide 16, i.e. less than 20 codons, a quadruplet code would be uneconomical in providing too many, i.e. 256 codon units. But how could molecules be so intelligent as to accept a logical argument? The answer is: It is the physical behaviour of the nucleotide sequences which determines the optimal choice among the different possible associations and the one chosen appears to be in agreement with our rational logic, because it offers evolutionary advantages with respect to precision and speed of information transfer. Precision requires interactions of sufficient stability, thus it involves "stickiness" which limits the rate of information transfer. The codon–anticodon interaction therefore must be optimised to involve:

1. sufficient functional capacities (i.e. requiring more than twenty combinations)
2. distinctive recognition (requiring at least base triplet interactions) and
3. sufficiently low stickiness (keeping the codon-unit as small as possible).

What I wanted to say is that all abstract qualities of nucleic acids which we associate with their function to store, transfer, and process information are reflected by certain unique physico-chemical properties. It is not sufficient to have just macromolecular species resembling a sequence of different digits. The "digits" in addition must provide very specific physical interactions to cause the inherent property of self-instruction and code formation. This quality, unique to the nucleic acids (and their interactions with proteins) provides the capacity of self-organisation according to—or in agreement with—our abstract principles of purposefulness, usefulness and rationality.

In this situation it is highly desirable to use any available experimental tool to enhance our knowledge about the "Physico-chemical Properties of Nucleic Acids". The three volumes which appear under this title offer a large repertoire of studies. Not all of them may be equally relevant for an understanding of the characteristics of information processing, which also involves highly specific interactions of nucleic acids with proteins. Nevertheless, all these studies will finally contribute to our basic understanding of those properties, which are behind the structural features, specific interactions, and dynamic performances of these unique macromolecules. In our age of molecular biology it may seem to be somewhat fashionable to do research in the field of nucleic acids. However, there will be a long persisting interest in this field before our knowledge about the molecular details will have brought about a complete understanding of the sophisticated organisation of the genome of a highly developed cell.

Spiegelman once characterised the central role of nucleic acids by saying jokingly: "The evolution of life is a trick of nature to ensure a faster and better reproduction of the nucleic acids".

November 1972 MANFRED EIGEN

Contents Volume 1

"Electrical, Optical and Magnetic Properties of Nucleic Acids and Components"

1. E. D. BERGMANN and HANNAH WEILER-FEILCHENFIELD: "The dipole moments of purines and pyrimidines".
2. D. VASILESCU: "Some electrical properties of nucleic acids and components".
3. M. A. SLIFKIN: "Charge transfer interactions of purines and pyrimidines".
4. M. DANIELS: "Recent developments in the fluorescence of DNA bases and DNA at 300°K".
5. C. HELENE: "Comparison of excited states and energy transfer in polynucleotides and aggregates of nucleic acid components".
6. C. NICOLAU: "Short-lived free radicals in aqueous solutions of nucleic acid components".
7. J. N. HERAK: "E.p.r. of irradiated single crystals of the nucleic acid constituents".
8. E. R. LOCHMANN and ASTRID MICHELER: "Binding of organic dyes to nucleic acids and the photodynamic effect".
9. A. VAN DE VORST AND Y. LION: "Photosensitisation of DNA's constituents by acridine dyes: an e.s.r. study".

AUTHOR INDEX
SUBJECT INDEX

Contents Volume 3

"Intra- and Intermolecular Interactions, Radiation Effects in DNA Cells, and Repair Mechanisms"

17. H. RICHARD: "Renaturation of DNA in the absence and in the presence of Cu^{2+} ions".
18. M. LENG, M. DOURLENT and C. HELENE: "Interactions in nucleic acids".
19. L. LERMAN: "The polymer and salt-induced condensation of DNA".
20. F. W. BENZ and G. C. K. ROBERTS: "The interaction of nucleotides with bovine pancreatic ribonuclease".
21. M. G. ORMEROD: "The measurement of radiation-induced strand breaks in the DNA of mammalian cells".
22. M. ERRERA and RAMAN: "Stability and evolution of DNA from the point of view of molecular radiobiology".

AUTHOR INDEX
SUBJECT INDEX

Contents Volume 2

Structural Studies on Nucleic Acids and Other Biopolymers

CONTRIBUTORS v
PREFACE vii
FOREWORD ix
CONTENTS VOL. 1 xv
CONTENTS VOL. 3 xv

10. Structural Studies of Nucleic Acid and Polynucleotides by Infrared and Raman Spectroscopy
K. A. HARTMAN, R. C. LORD AND G. J. THOMAS JR.

I. Introduction 1
II. Infrared studies of nucleic acid structure 13
III. Raman spectroscopy and nucleic acid structure 46
References 87

11. Infrared and Raman Spectra of Nucleic Acids—Vibrations in the Base Residues
M. TSUBOI, S. TAKAHASI AND I. HARADA

I. Introduction 92
II. The spectra of the four-base residues 93
III. Data for characterising Raman lines and infrared absorption bands . 98
IV. Normal coordinate treatment 103
V. Characterisation of each Raman line or infrared band 111
VI. Effects of inter-base interactions 133
Acknowledgements 143
References 143

12. Conformation of Nucleic Acids, Oligo- and Polynucleotides by Circular Dichroism Investigations
C. A. BUSH AND J. BRAHMS

I. Introduction 147
II. Mononucleosides 148
III. Analysis of polynucleotide CD 162
IV. Nucleic acids 178
References 183

13. Nuclear Magnetic Resonance Studies of Transfer RNAs in Solution
M. P. SCHWEIZER, S. I. CHAN AND J. E. CRAWFORD

I. Introduction 187
II. The interaction of purine with *E. coli* $tRNA_f^{Met}$ 190

III. Organic solvent denaturation of transfer RNA 195
IV. Summary 208
Acknowledgements 209
References 209

14. The Structure of Viral RNA
W. FIERS

I. Viral RNA genomes 213
II. Chemical identification of endgroups 216
III. The primary structure of viral RNA 220
IV. Sequences of the viral RNA not involved in translation . . . 221
V. Sequences of the viral RNA which contain the message for polypeptide synthesis 224
VI. Physical structure of the viral RNA 226
Note added in proof 232
Acknowledgements 232
References 232

15. Thermodynamics and Kinetics of Conformational Transitions in Oligonucleotides and *t*RNA
D. RIESNER AND R. RÖMER

I. Introduction 237
II. Single strand stacking 238
III. Base pairing in oligonucleotides and *t*RNA fragments 244
IV. Conformational changes in *t*RNA 273
V. Biological implications 309
Acknowledgements 312
Addendum 312
References 314

16. Hydrodynamic Properties and Conformational Changes of Linear Native DNA
K. E. REINERT

I. Introduction 320
II. Theoretical aspects of DNA hydrodynamics 321
III. Special experimental aspects of DNA hydrodynamics and molecular weight determination 330
IV. Hydrodynamic behaviour and DNA molecular weight 335
V. DNA as a polyelectrolyte 341
VI. Analysis of induced DNA conformational changes 347
Acknowledgements 353
References 353

AUTHOR INDEX 357
SUBJECT INDEX 371

CHAPTER 10

Structural Studies of Nucleic Acids and Polynucleotides by Infrared and Raman Spectroscopy

K. A. HARTMAN

Department of Microbiology and Biophysics, University of Rhode Island, Kingston, Rhode Island, U.S.A.

R. C. LORD

Department of Chemistry and Spectroscopy Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.

and G. J. THOMAS, JR.

Department of Chemistry, Southeastern Massachusetts University, North Dartmouth, Massachusetts, U.S.A.

I. Introduction 2
A. Nature of the data of vibrational spectroscopy 2
B. Brief survey of experimental infrared and Raman spectroscopy . 5
C. Origins of vibrational spectra 8
D. Some factors affecting characteristic frequencies of nucleic acid constituents 10

II. Infrared Studies of Nucleic Acid Structure 13
A. Structure and interactions of nucleic acids and their constituents . 13
B. Hydration and structure of nucleic acids 18
C. Conformation of RNA in aqueous solution 33
D. Specificity of the hydrogen bonding of purine and pyrimidine derivatives in deuterochloroform solution 39

III. Raman Spectroscopy and Nucleic Acid Structure 46
A. Structures and interactions of nucleic acid constituents . . . 46
B. Conformational studies of polynucleotides and nucleic acids . . 69
References 87

I. Introduction

A. NATURE OF THE DATA OF VIBRATIONAL SPECTROSCOPY

Molecular biology deals with the relationship between molecular structure and biological function. It advanced tremendously when the methods of X-ray diffraction were successfully applied to a number of biological polymers to yield the detailed molecular structures of these materials. With precise information about the structure of an enzyme, for example, its functioning as a catalyst of biochemical reaction can be elucidated in more definite fashion than can be done in the absence of such information. However, there are two practical limitations to the X-ray crystallographic approach: firstly, it is a time-consuming and complicated procedure and secondly, it can be applied only to crystallisable systems. Nucleic acids *in vivo* exist in dilute aqueous solution, where structural differences from the crystalline state may be significant. Thus it is important to have other physical techniques capable of giving structural information, particularly in aqueous solution, and vibrational spectroscopy is one of the most promising of these.

It has been recognised for nearly a century that intramolecular vibrations produce absorption in the infrared region of the spectrum and that the observed absorption frequencies are characteristic of the small groups of atoms known to the organic chemist as functional groups. Despite the long history of this evidence of a relationship between structural chemistry and infrared absorption, the difficulty of working in the infrared region in the pre-electronic era delayed the general use of infrared spectroscopy by the organic chemist until after World War II. In fact, the discovery of the Raman effect in 1928 stimulated more work on the correlation of molecular structure with vibrational frequencies than had previously accumulated from infrared absorption spectroscopy. This was due to the relative ease of photographic Raman spectroscopy of that era and to the fact that the kind of information furnished by the Raman effect is essentially the same as that by infrared absorption.

At present the development of instrumentation for routine determination of both infrared and Raman spectra has proceeded to the point where one can use either technique or both to aid in the solution of problems in structural chemistry. The kind of information provided by vibrational spectroscopy, whether infrared or Raman, is:

(1) The frequencies ν of intramolecular and intermolecular vibrations in the range 0·3–120 THz (1 THz$=10^{12}$ cycles per sec). The more customary units are wavenumber σ in cm^{-1}, in which unit the above range is 10–4000 cm^{-1} and wavelength λ in micrometers, 1000–2·5 μm.

(2) The intensities in appropriate units of the spectral lines or bands associated with these frequencies. It is customary to plot percent trans-

mission T_σ as a function of wavenumber in the infrared, and this quantity is related to a molar property, the absorptivity a_σ at wavenumber σ, by the definition

$$a_\sigma = (1/bc) \log_{10} (100/T_\sigma) \quad (1)$$

in which b is the length in cm of the absorbing path of the radiation through the sample and c the sample concentration in mol/cm^3. In the Raman effect it is difficult to obtain a molar property analogous to a_σ, and it is customary to plot the intensity of the Raman spectrum on a relative scale. For quantitative purposes some arbitrary standard of scattered intensity (such as that of a given line in the spectrum of carbon tetrachloride, for example) may be used as a reference to which the desired spectrum can be compared under identical conditions of excitation. Typical infrared and Raman spectra recorded instrumentally to the same wavenumber scale are given in Fig. 1.

(3) Polarisation characteristics of the absorbed infrared radiation or scattered Raman radiation. If a solid sample has its molecules so ordered that the macroscopic optical properties of the sample are different in different directions, the infrared spectra taken with radiation polarised parallel to these different directions will be different (see Section B, below). In particular, if crystalline symmetry dictates that the dipole-moment change associated with a given molecular vibration be parallel to a given direction in a crystalline sample, radiation polarised parallel to such direction will be absorbed whereas that polarised perpendicular will not. Thus the polarisation characteristics can give information about the direction of the vibrating dipole moment in an oriented sample and, under favourable circumstances, about the orientation of functional groups in the sample (Section IIB).

The nature of polarised Raman scattering also depends on the extent to which the sample is oriented. If the sample is monocrystalline, which is not often the case for biological macromolecules, much information about the orientation of functional groups can be obtained from polarised Raman studies (see, for example, the treatment of Fanconi *et al.*, 1969). Even when the sample possesses no macroscopic orientation, characterisation of a functional group by the depolarisation ratio ρ of one of its Raman frequencies may be possible. For scattering at 90° to the incident beam, the depolarisation ratio is defined as

$$\rho = I_\perp / I_\parallel \quad (2)$$

where $I_\perp$ and $I_\parallel$ are the intensities of the Raman radiation polarised respectively perpendicular and parallel to the scattering plane, that is, the plane normal to the incident beam. If the incident radiation is plane polarised perpendicular to the direction in which the scattered radiation is collected, ρ lies in the range 0–$\frac{3}{4}$; for unpolarised incident radiation, the range is 0–$\frac{6}{7}$.

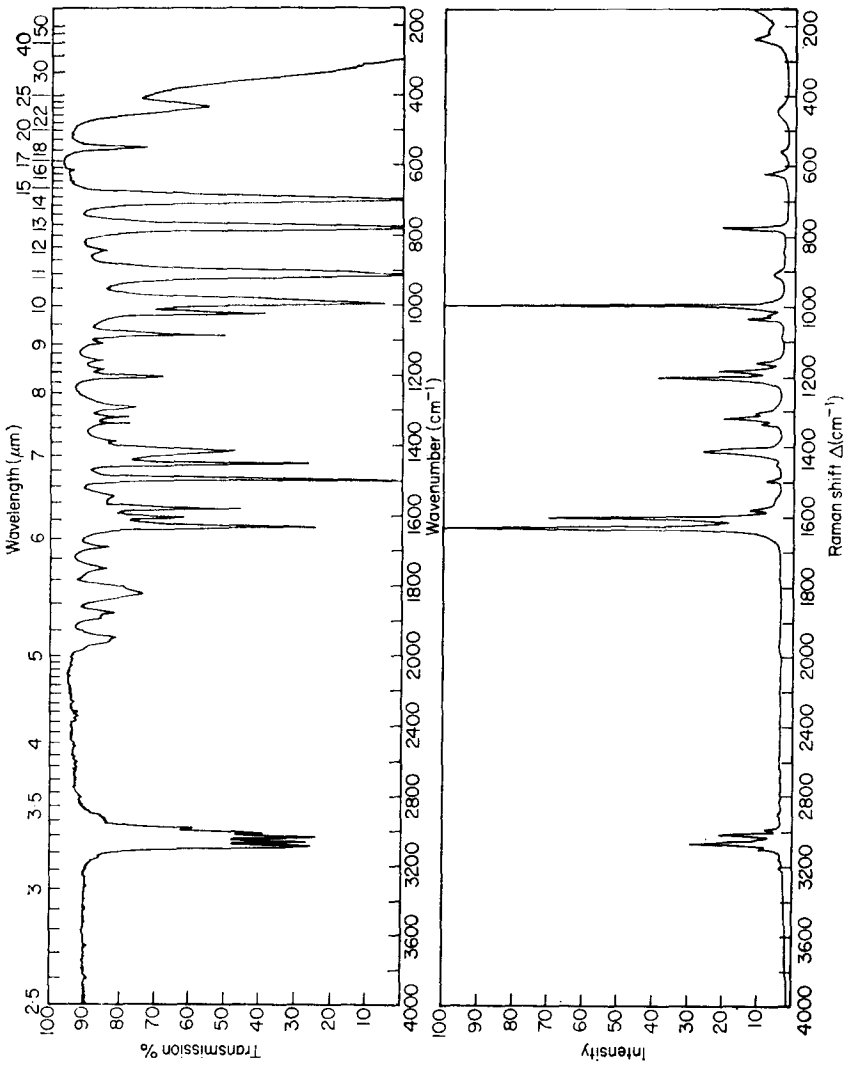

FIG. 1. Infrared spectrum (upper) and Raman spectrum (lower) of liquid styrene. Transmission minima of infrared bands are usually found at the same wavenumbers as the intensity maxima of Raman lines, but relative intensities in the two spectra differ greatly. Thus the data provided by the two spectra are complementary.

B. BRIEF SURVEY OF EXPERIMENTAL INFRARED AND RAMAN SPECTROSCOPY

We consider here briefly the general requirements which must be met by the sample in order to obtain infrared and Raman spectra of acceptable quality. More detailed treatment of these topics and of instrumentation for both techniques may be found in recent review articles (Thomas, 1971; Tsuboi, 1972).

1. *Infrared spectroscopy*

An important practical advantage of infrared spectroscopy as an analytical technique is that it can be applied to most materials regardless of such gross physical and chemical properties of the sample as phase, colour, turbidity, molecular weight and the like. Ordinarily, nucleic acids and their constituents are studied as solids (oriented films or fibres, Section IIB, below) or liquid solutions (aqueous and chloroform, Sections IIA, C and D, below).

a. *Films.* Films are prepared by evaporating an aqueous solution of the sample on a plate of infrared window material (AgCl, CaF_2 or BaF_2) which can then be mounted directly in the spectrometer. Usually about 0·5 mg of the sample is dissolved in water, spread over an area of about 100 mm^2 and dried so that a film of suitable thickness (~ 10 μm) is obtained. Since the molecular conformation of a nucleic acid often depends on the humidity of the surrounding air, films of nucleic acids are ordinarily maintained in sealed cells in which a salt solution is placed to achieve the desired humidity (Falk *et al.*, 1962). Samples may also be deuterated in this manner. For example, if a saturated solution of sodium bromate in D_2O is placed in the bottom of the cell, virtually all OH and NH groups of the sample are exchanged by deuterium.

The structural information derivable from the infrared spectrum of a nucleic acid film may be increased if the molecules in the film can first be oriented in a preferred direction and then the spectrum of the film determined with polarised infrared radiation. Oriented films of nucleic acids and polynucleotides are made by stroking the wet fibres unidirectionally until dry. The oriented film is then mounted so that the molecules have one specific orientation (say parallel) with respect to the plane of polarisation of the radiation and its spectrum is recorded. The orientation of the sample is then changed by 90° and a second spectrum is obtained. If a particular absorption band shows different absorbances in the two orientations it is said to be "dichroic" (see Section IIB, below).

b. *Aqueous solutions.* Because of the opacity of water to infrared radiation, special techniques are required to obtain satisfactory spectra of aqueous

solutions (Thomas, 1971). Nucleic acids and polynucleotides are usually studied in D_2O solutions in the region 1450–1750 cm^{-1}, where double-bond vibrations of the bases give rise to intense infrared bands and the interfering absorption by the solvent is easily compensated. Cells for aqueous solution spectroscopy may be constructed of fluorite (CaF_2) windows. Ordinarily a cell thickness in the range 25–75 μm is most satisfactory when the solute concentration is in the range 100–30 mg/ml (10–3% by weight). A minimum volume of about 7–20 μl is required to fill the cell. Techniques for more dilute solutions have been described by Miles (1968). The use of acetate or citrate buffers and the exposure of solutions to atmospheric water vapour should be avoided since these cause interfering absorption in the double-bond region.

c. *Chloroform solutions.* Infrared studies of chloroform solutions of nucleoside derivatives provide important information on the specificity of base pairing and strength of hydrogen-bonding interaction (Section IID, below). An essential requirement in such studies is that the solvent and solute be free of impurities or additives which may compete with hydrogen bonding between solute molecules. For example, reagent grade chloroform contains a preservative (usually ethanol) which must be removed, before use, by passage over a column of dry alumina gel. Studies of chloroform solutions have usually been carried out in the NH stretching region, 3300–3500 cm^{-1}, with cells of fused silica. The region of transparency to infrared radiation may be further increased by the use of deuterochloroform ($CDCl_3$) as the solvent (Thomas, 1971).

2. *Raman spectroscopy*

Chemical purity and homogeneity of the sample must be more carefully controlled for Raman spectroscopy than for infrared spectroscopy. Physical properties of the sample, such as colour, optical homogeneity and the like, must also be taken into account in preparing a sample so that its Raman spectrum can be recorded. In other respects, sample-handling for Raman spectroscopy is simpler than for the infrared; for instance, cells may be made of glass, which greatly simplifying the handling of aqueous samples.

Nucleic acids and polynucleotides are most frequently investigated as aqueous solutions (Sections IIIA and B, below) and sample-handling procedures for these are briefly discussed.

a. *General requirements of the sample.* In order to obtain a satisfactory Raman spectrum it is necessary that interaction of the radiation with the sample by such other mechanisms as absorption, fluorescence and Tyndall scattering be minimised.