

**RAMAN OPTICAL ACTIVITY  
OF BIOLOGICAL MOLECULES**

**A thesis submitted  
for the partial fulfillment of the degree of**

**Doctor of Philosophy**

**in the Chemistry Department,  
the Faculty of Science,  
The University, Glasgow**

**by**

***Zai-Qing Wen***

**October 1992**

ProQuest Number: 11007745

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 11007745

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 – 1346

GLASGOW  
UNIVERSITY  
LIBRARY

*Thesis*  
*9350*  
*copy 1*

**To my mother,  
my deceased father,  
and my family.**

## Acknowledgements

I wish to express my thanks to Professor Laurence D. Barron, my supervisor, who is not only an enthusiastic giant scientist, but also a warm-hearted human being, for his extraordinary patience, proficient advice and persistent encouragement during the course of this study.

I would like to thank my family, who bear with me during the study. I am particularly indebted to my daughter for leaving her three years.

Many thanks are given to my colleagues in the Chemistry Department for their friendliness and assistance, particularly to Dr. L. Hecht, Dr. A. R. Gargaro, Mr. A. F. Bell and Mr. S. J. Ford, with whom we had many discussions. My thanks are also extended to staff in the Chemistry Department, their names are too many to be listed, for their kind assistance, which made my stay in Glasgow enjoyable and memorable.

Finally, I would like to thank my friends and colleagues in China with whom we had many communications which encouraged and supported my study here.

## Abbreviations

BSA	Bovine Serum Albumin
CCD	Charge-Coupled Device
CD	Circular Dichroism
$\alpha$ -CH	$\alpha$ -Chymotrypsin
CID	Circular Intensity Difference
DCP	Dual Circular Polarization
ECD	Electronic Circular Dichroism
HBD	Holographic Bragg Diffraction
ICP	Incident Circular Polarization
INS	Insulin
$\beta$ -GL	$\beta$ -Lactoglobulin
LYZ	Lysozyme
$\alpha$ -LC	$\alpha$ -Lactalbumin
ORD	Optical Rotatory Dispersion
OVA	Ovalbumin
RNS	Ribonuclease A
ROA	Raman Optical Activity
SCF	Self-Consistent Field
SCP	Scattering Circular Polarization
SNR	Signal to Noise Ratio
UVCD	Ultra-Violet Circular Dichroism
VCD	Vibrational Circular Dichroism
VOA	Vibrational Optical Activity

# Contents

ABSTRACT	1
<b>CHAPTER 1. Introduction</b>	<b>4</b>
1.1. Optical activity	4
1.2. Vibrational optical activity	6
1.2.1. Vibrational circular dichroism	9
1.2.2. Vibrational Raman optical activity	11
<b>CHAPTER 2. Basic theory of Raman optical activity</b>	<b>20</b>
2.1. Introduction	20
2.2. Molecular multipole moments	21
2.3. Light scattering by a molecule	22
2.4. Basic Raman optical activity equation	26
2.5. Molecular scattering transition tensors	28
2.6. The bond polarizability model of ROA and the merits of backscattering	30
2.7. The <i>ab initio</i> ROA calculation	32
<b>CHAPTER 3. Raman optical activity instrument</b>	<b>38</b>
3.1. Introduction	38

3.2	Some basic considerations for ROA instrumentation	40
3.3	The optical configuration of the ROA instrument	42
3.3.1	Light source	44
3.3.2	Polarizer	45
3.3.3	Light modulator	45
3.3.4	Focusing lens	46
3.3.5	Backscattering block	47
3.3.6	Sample cell	48
3.3.7	Collection optic and magnification	49
3.3.8	Spectrograph	50
3.3.9	Holographic Bragg diffraction filter	51
3.3.10	CCD detector	51
3.4	Electronics and computer system	54
3.4.1	CCD electronics	54
3.4.2	Polarization modulation control	56
3.4.3	Computer and software	56
3.4.4	ROA spectrum acquisition	57
3.5	Spectrum calibration	58
3.6	Experimental details	59
3.6.1	Sample handling	59
3.6.2	Artifact control	60
3.6.3	Stray light and base line	62
3.7	The performance of the ROA instrument	63
3.7.1	Instrument sensitivity	63
3.7.2	Spectral resolution	69
3.7.3	Spectral coverage	69

<b>CHAPTER 4. Raman optical activity of peptides and polypeptides</b>	<b>72</b>
4.1 Introduction	72
4.2 Raman optical activity of di-L-alanine	74
4.3 Raman optical activity of tri-, and tetra-L-alanine	92
4.4 Raman optical activity of L-Pro-L-Leu-Gly-amide	97
4.5 Raman optical activity of poly-L-lysine and poly-L-glutamic acid	103
4.6 Conclusions	109
<b>CHAPTER 5. Raman optical activity of proteins</b>	<b>113</b>
5.1 Introduction	113
5.2 Experimental section	116
5.3 Results and discussions	118
5.3.1 Bovine serum albumin, insulin and ovalbumin	120
5.3.2 Lysozyme and $\alpha$ -lactalbumin	126
5.3.3 Ribonuclease A, $\alpha$ -chymotrypsin and $\beta$ -lactoglobulin	131
5.3.4 $\alpha$ -lactalbumin at low pH and upon calcium depletion	136
5.4 Conclusions	141
<b>CHAPTER 6. Raman optical activity of carbohydrates</b>	<b>147</b>
6.1 Introduction	147
6.2 Experimental section	150
6.3 Results and discussions	151

6.3.1	Absolute configuration	153
6.3.2	Anomeric configuration	153
6.3.3	Homomorphic sugars	161
6.3.4	Epimeric sugars	168
6.3.5	Conformation of the CH <sub>2</sub> OH group	171
6.3.6	Furanose and pyranose	173
6.3.7	The glycosidic linkage and the merit of the ROA $\Delta$ value	176
6.4	Conclusions	179
PUBLICATIONS		184

## Abstract

This thesis describes work that has helped to establish Raman optical activity (ROA) as a powerful new chiroptical spectroscopic technique for the study of molecular chirality and conformation of biological molecules in aqueous solution.

The first chapter describes the background and recent developments of vibrational optical activity including both the infrared vibrational circular dichroism (VCD) and Raman optical activity approaches.

In chapter two, the basic theory of vibrational Raman optical activity is briefly reviewed. It comprises the fundamental theory to describe the vibrational Raman optical activity phenomenon, the rationale of the choice of the backscattering geometry for ROA instrument set-up and the basis of the more advanced *ab initio* ROA theory for calculation of ROA spectra.

A new ROA instrument based on the backscattering geometry and back thinned CCD (Charge-Coupled Device) light detector is detailed in chapter 3. The new ROA instrument represents the contemporary development of ROA instrumentation and the up-to-date sophisticated optical and electronic devices used in the ROA spectrometer. A few basic considerations in ROA instrumentation and the performance of the instrument are discussed. The breakthrough of the instrument sensitivity has enabled ROA spectroscopy to be applied to important biological molecules in aqueous solution for the first time. The following three chapters are devoted to the ROA study of a number of biological

molecules including small peptides, polypeptides, proteins and carbohydrates. These ROA data constitute the basis of ROA study for more complicated biological molecules in the future.

Chapter 4 deals with ROA studies on model peptides and polypeptides. They comprise L-alanine oligomers: di-L-alanine and tri-, tetra-L-alanine, and a tripeptide L-Pro-L-Leu-Gly-amide for model  $\beta$ -turn structure plus two polyamino acids poly-L-glutamic acid and poly-L-lysine. Di-L-alanine and its enantiomer are investigated in detail in various aqueous solutions and the results suggest that di-L-alanine could be used as a good model peptide for vibrational optical activity spectra analysis of peptides and polypeptides. ROA is a very local effect that is related to the intrinsic chirality. The most important ROA bands of peptides are in the extended amide III region.

In chapter 5, the first ROA spectra of eight globular proteins in aqueous solution are reported and a preliminary empirical analysis presented. These protein ROA data clearly demonstrate that ROA is now able to investigate protein structure in the solution phase. The dominant ROA features of proteins arise mainly from the polypeptide backbone. Proteins containing different secondary structure compositions show characteristic ROA patterns. The most prominent ROA features are concentrated in the extended amide III region and are particularly sensitive to reverse turn structure. The ROA band intensity and  $\Delta$  value offer a sensitive probe of the rigidity or flexibility of the globular proteins.

In the last chapter 6, the ROA spectra of a range of carbohydrates including fifteen monosaccharides, a disaccharide and a cyclodextrin are

investigated. The overwhelming ROA spectral information convincingly demonstrates that carbohydrates are particularly favourable samples for vibrational ROA study. ROA measurements on carbohydrates can yield ample stereochemical information with respect to the glycosidic linkage, anomeric configuration, sugar ring chair conformation and intramolecular interaction between adjacent chiral centres. Of all the information available from carbohydrate ROA spectra, the characteristic ROA couplet of the glycosidic linkage is probably the most valuable, which may be used to probe the conformation of disaccharides, oligosaccharides and polysaccharides.

## CHAPTER 1

# Introduction

### 1.1. Optical activity

Optical activity phenomena originate in the different response of chiral molecules to left and right circularly polarized light.<sup>1,2</sup> The origin of optical activity is that a chiral molecular structure is not superimposable on its mirror image. Optical activity is very sensitive to the spatial disposition of atoms in a molecule and is therefore regarded as a molecular property directly associated with molecular geometry. The study of optical activity in organic molecules had directly led to the establishing of the chemical concept that molecules have three dimensional structure in last century.<sup>1</sup> Measurement of optical activity has played a crucial role to determine relative configuration and conformations in stereochemistry.<sup>3</sup>

The significance of optical activity measurement of biological molecules is born out by the simple fact that most biologically important molecules, ranging from amino acids, peptides, proteins, to nucleic acids and carbohydrates are chiral.<sup>4</sup> Proteins are usually made up of L-amino acids while the walls of living cells, and nucleic acids, contain the D-enantiomers of sugars. Most enzymes react with only one enantiomer of a substrate and so have chiral selectivity. The optical activity of molecules play a vital role in pharmaceutical chemistry. As is well known, over 50% of commercial drugs are chiral compounds. Two enantiomers usually have

different pharmacological response in that one enantiomer is an effective drug while the other is useless or even harmful. The tragedy caused by the adverse effect of the S-enantiomer of thalidamide, which was used as a sedative and sleeping drug in the 1960s and caused serious malformations in babies born to mothers who had taken the drug in early pregnancy, is one particularly notorious example.<sup>5</sup>

Conventionally, optical activity was mainly measured through two spectroscopic techniques, namely optical rotatory dispersion (ORD) and electronic circular dichroism (ECD).<sup>3</sup> The former refers to the measurement of the wavelength dependence of the rotation of plane-polarized light by a chiral medium and the latter measures differential electronic absorption spectra of chiral molecules under left and right circularly polarized light irradiation. They are usually called chiroptical spectroscopic methods because they are sensitive to molecular chirality, thereby being distinguished from general spectroscopic techniques. Spectra of ORD and ECD are inherently related to each other through a mathematical relationship known as Kronig-Kramer transform and give rise to the same information. An ORD spectrum thus can be calculated through the transform from the corresponding ECD spectrum, or vice versa. ORD was developed first for its relatively simple measurement technique, but was superseded by ECD with the advance of sophisticated optic-electronic technology. Moreover, a ECD spectrum has generally better resolution than that of ORD.

Optical activity measurements on biological molecules through ORD and ECD has contributed a great deal of information on biological

molecules in the solution phase in biochemistry since the middle of this century. ECD has been employed to determine the compositions of secondary chain conformation of proteins in solution<sup>6</sup> which is not amenable by the X-ray diffraction. An excellent example is the prediction of the left hand DNA helix from ECD study in solution in advance of the X-ray results.<sup>7</sup> Taking advantage of the high sensitivity, time resolved ECD can probe the dynamics of biological macromolecules.<sup>8</sup> It has also been used to probe the side group structures of proteins and protein-ligand interaction. In the theoretical aspect, the interpretation of ECD spectra has been placed on a firm quantum mechanical basis and calculation of ECD spectra can be carried out even on complex systems such as biopolymers.<sup>9</sup> It is now regarded as a standard technique in biomolecular structure study. However, limitations are inherently associated with these conventional optical activity measurement techniques. One of the disadvantages is the narrow spectral region accessible to ECD, which is usually limited to the UV-visible region although vacuum techniques enable ECD to extend into the far ultraviolet region. Another drawback is the poor resolution of a ECD spectrum in that bands are generally overlapped.

## **1.2. Vibrational optical activity**

In contrast to traditional optical activity studies, vibrational optical activity is a rather young field born only two decades ago. It probes optical activity through vibrational transitions while the conventional optical activity is related to electronic transitions. It includes two independent but complementary approaches, namely vibrational circular

dichroism (VCD)<sup>10</sup> and Raman optical activity (ROA).<sup>2</sup> VCD is the natural development of circular dichroism from the UV-visible to the infrared region and measures the differential absorption infrared spectra of chiral molecules in circularly polarized infrared radiation. ROA probes vibrational optical activity through light scattering and measures a slight differential Raman scattered intensity of a chiral substance using right and left circularly polarized incident light (Fig. 1.1).

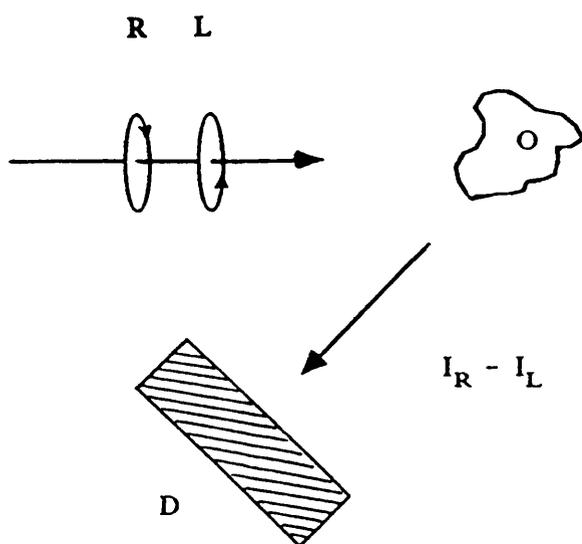


Fig. 1.1 The basic ROA experiment. The observable is  $I_R - I_L$ , where  $I_R$  and  $I_L$  are the scattered Raman intensities of the molecule in right (R) and left (L) circularly polarized incident light. O is the molecular origin and D is the detector.

The merit of vibrational optical activity studies is illustrated by the fact that they investigate molecular chirality through vibrational transitions on the electronic ground state. This renders a much wider spectral range (the entire vibrational spectrum) accessible to optical activity measurement. Vibrational spectra generally contain many well-resolved bands and are relatively easily assigned, for example, to specific function groups. Fundamentally, a vibrational spectrum is the manifestation of vibrational modes associated with every atoms of the molecule. By contrast, electronic absorption spectra is only related to the chromophores, the numbers of which in a molecule are rather limited. Furthermore, The use of infrared and visible radiation as exciting light in vibrational optical activity measurement is beneficial to the study of biological molecule because ultraviolet radiation, especially the vacuum

Table 1-1. Characteristics of three major chiroptical techniques.

	ROA	VCD	ECD
radiation	visible	infrared	UV-visible
modulation	optical-electric	photoelastic	optical-electric
spectral coverage	3500-50cm <sup>-1</sup>	4000-600cm <sup>-1</sup>	700-180nm
resolution	high	high	low
sensitivity	$\Delta I/I \sim 10^{-5}$	$\Delta A/A \sim 10^{-5}$	$\Delta A/A \sim 10^{-7}$
physical origin	vibrational	vibrational	electronic
instrumentation	developing	near-commercial	commercial
suitable solvent	H <sub>2</sub> O, D <sub>2</sub> O, organic	D <sub>2</sub> O, organic	H <sub>2</sub> O, D <sub>2</sub> O
damage to sample	no	no	yes ( in UV )

ultra-violet light used in vacuum UV-CD, is harmful to many biological molecules and may cause cleavage of the chemical bonds which can be avoided using infrared and visible radiation.

Table 1 lists the comparison of the three major chiroptical spectroscopic techniques in terms of their experimental characteristics and applicability in studies of biological molecules.

### 1.2.1 Vibrational circular dichroism

The first genuine VCD spectrum was obtained on optically active crystals  $\alpha$ -NiSO<sub>4</sub>·6H<sub>2</sub>O by Hsu and Holzwarth in 1973.<sup>10</sup> Subsequently, solution phase VCD in the C-H stretching region was reported in 1974 by Holzwarth et al.<sup>11</sup> followed by substantial VCD data in the C-H, O-H, N-H region after the improvement of instrument sensitivity by Nafie et al.<sup>12</sup> These experimental results opened the new field to investigate optical activity *via* infrared radiation. However, measurement of VCD spectra offers a great challenge to vibrational spectroscopist because VCD is a very weak effect that is 3 to 5 orders of magnitudes smaller than the parent infrared spectra and the infrared modulation and detection techniques were under-developed compared to in the visible region. During the last two decades, much effort has been devoted to the VCD instrumentation.<sup>13,14</sup> VCD spectrometers are now nearly commercially available and high quality VCD spectra of organic chiral molecules can be routinely recorded in the spectral range between 4000-600cm<sup>-1</sup>.<sup>14</sup>

Two principal infrared spectrometers are employed in VCD

measurements. One employs the traditional dispersive scanning IR spectrometer plus a photoelastic modulator and the other uses the more sophisticated FTIR spectrometer incorporated with a modulator. The dispersive scanning scheme has a better signal to noise ratio than FTIR in certain regions. Its detection limit can be as low as  $\Delta A/A \sim 5 \times 10^{-6}$  in favourable samples. On the other hand, FTIR-VCD has better resolution and it is nearly commercially available. Since solvent absorption is a severe problem in the entire infrared region, one usually can only obtain VCD in a limited spectra region in which case dispersive VCD would be favoured. If the entire spectrum and better resolution are needed, FTIR-VCD would be certainly preferred. Therefore, dispersive scanning and FTIR-VCD will complement each other in VCD spectra measurement.

In the theoretical aspect, a large number of models have been proposed to predict VCD spectra.<sup>15</sup> They calculate the vibrational absorption spectra and circular dichroism spectra through the molecular properties  $\epsilon$ , the molar extinction coefficient, and  $\Delta\epsilon$ , the differential molar extinction coefficient which in turn are proportional to the dipole strength  $D = \langle m | \mu_\alpha | n \rangle \langle n | \mu_\alpha | m \rangle$  and rotational strength  $R = \text{Im}[\langle m | \mu_\alpha | n \rangle \langle n | m_\alpha | m \rangle]$  of a fundamental vibrational transition from the ground state  $m$  to the excited one  $n$ , where  $\mu_\alpha$ ,  $\mu_m$  are the electric and magnetic dipole moment operators respectively. Although all these simple models have been successful to some extent in understanding the mechanisms generating VCD, none of them are of general application. The most promising one is probably the *ab initio* quantum mechanical calculation which now can calculate VCD spectra of small chiral molecules with fairly good agreement with data obtained by experiment.<sup>16, 17</sup>

The study of VCD was initially concentrated on favourable organic compounds to test theoretical models. With the improvement of the instrument sensitivity, applications to biological model compounds were initiated.<sup>14</sup> Extensive studies on model biopolymers and typical proteins have been carried out. Chain length dependence studies of VCD on a number of given length oligopeptides has shown that VCD has a very local structural sensitivity that is different from conventional ECD. It has been demonstrated that VCD bands in the amide I and amide II region are sensitive to secondary chain conformation of polypeptides and proteins<sup>18</sup>. Characteristic VCD patterns in the amide I region have been correlated to the corresponding protein secondary structure arrangement and statistical evaluation of the compositions of protein secondary structure have been performed. Taking advantage of good instrument sensitivity, temperature and pH dependence studies of VCD have also been conducted on model polypeptides and proteins.<sup>19</sup> One pitfall of VCD inherited from infrared absorption spectra is that water is not a suitable solvent for VCD measurement of biological molecules because water absorbs strongly infrared radiation and D<sub>2</sub>O has to be employed, which usually changes the molecular vibrational transitions associated with certain labile protons such as in N-H and O-H groups.

### **1.2.2 Vibrational Raman optical activity**

Vibrational Raman optical activity was first predicted by Atkins and Barron two decades ago during the study of polarized light scattering of molecules.<sup>20</sup> The basic Raman circular intensity difference equation was

established by Barron and Buckingham<sup>21</sup> based on semi-classical treatment and, with Bogaard, they observed the first genuine ROA spectrum in 1973.<sup>22</sup> These studies offered an alternative approach to investigate vibrational optical activity through light scattering, which is completely different from the more conventional measurement of vibrational optical activity through absorption as in VCD.

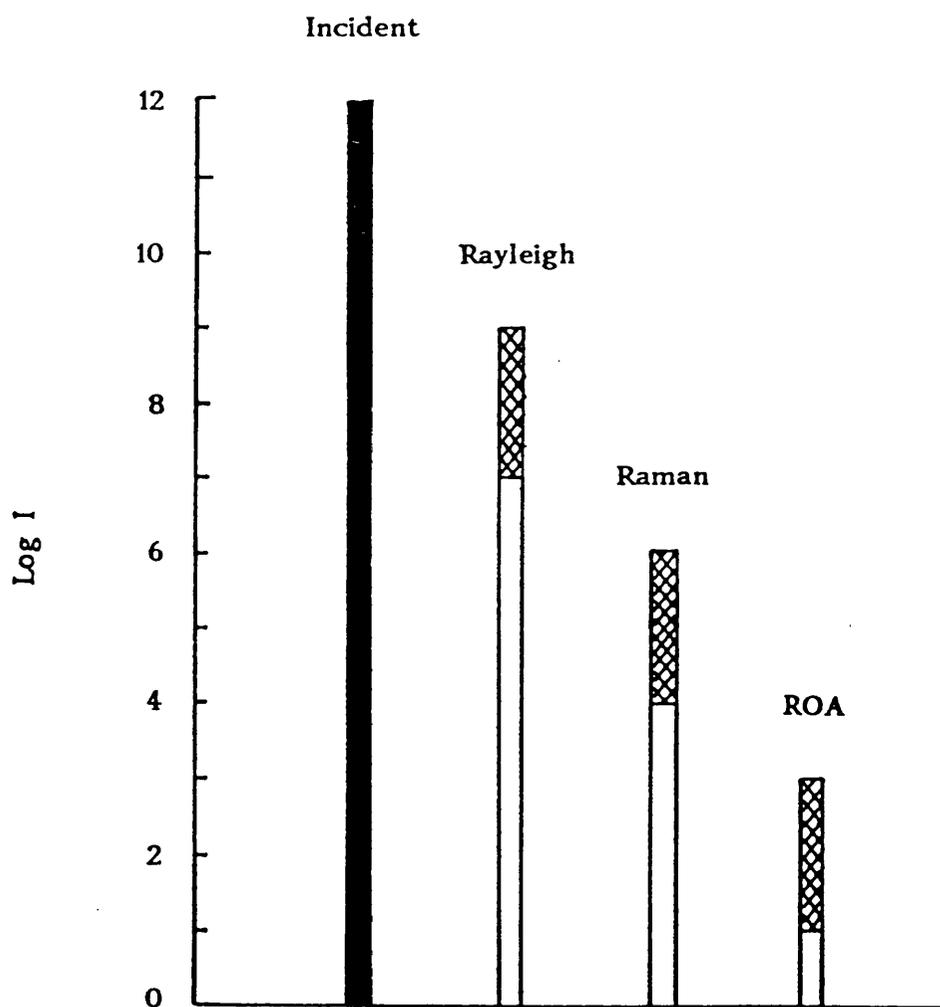


Fig. 1.2 The relative intensity of three major light scattering phenomena.

The development of Raman optical activity, however, encountered a number of obstacles with respect to both experimental and theoretical aspects for the fundamental reason that ROA is a second order perturbation phenomenon from the standpoint of quantum mechanics. The ROA effect is so weak that only  $\sim 10^{-3}$ - $10^{-4}$  Raman scattered photons show the circular polarization difference noted for each of the single Raman photons typically generated by an incident beam flux of  $10^8$  photons. Fig. 1.2 shows the relative intensity of the three basic light scattering phenomena under a normal incident light. It used to take a very long time (days) to record a reasonable entire ROA spectrum on favourable organic neat liquids in the early days of ROA study.

Moreover, artifacts were another barrier to routine measurement of ROA and only depolarized ROA spectra could be recorded without severe artifact contamination.<sup>22</sup> The elimination of artifacts was not realized until a detailed quantitative analysis of the mechanism of artifacts was undertaken and corresponding procedures to suppress it were introduced.<sup>23-25</sup> Another obstacle to ROA experiments, which is even more severe in the modern multichannel Raman spectrometers, is the Rayleigh stray light associated with Raman scattering which prevents the reliable recording of low frequency ROA spectra which probably contains more important information on molecular skeleton.

In the theoretical aspect, one also faces great challenge. The theory of ROA is difficult for it deals with the interaction of electromagnetic radiation with second order polarizability and optical activity tensors and involves the theories of both the electronic and vibrational states of low

symmetry molecules. Most of the earlier ROA studies were therefore restricted to fundamental research and favourable small organic chiral molecules<sup>26, 27</sup> despite the fact that it should be in biomolecules that ROA finds most application.

Breakthroughs in both experimental and theoretical ROA studies emerged in the last few years. The adoption of a backscattering geometry<sup>28</sup> for ROA measurements marked a cornerstone for two reasons. First, the measurement speed of an ROA spectrum is increased nearly eightfold in backscattering compared with the right angle arrangement which was previously employed in all ROA instruments. Second, the artifacts that had plagued ROA experiments from the very beginning were almost completely suppressed to below the noise level. Advances in optical-electronic technology offered an almost perfect light detector in the form of a charge-coupled device that ROA has long awaited.<sup>29</sup> Its high quantum efficiency, wide dynamic range and low readout noise boost the signal to noise ratio of the ROA instrument. Furthermore, new optical technology provided two critical optical components: a holographic notch filter<sup>30</sup> which rejects stray light at low frequency and a high-efficiency single-grating spectrograph. With the implementation of all these high technologies, our new generation ROA instrument<sup>31</sup> acquired a 100 fold or more increase in measurement speed, which enabled ROA measurements to be made for the first time on biological molecules in aqueous solution.

Other advances in the ROA experiment include the use of different scattering geometries based on the angle dependence of ROA observables. In  $90^\circ$  scattering, a magic angle was identified<sup>32</sup> in which the analyser is

set at  $54.74^\circ$  to the scattering plane and pure electric dipole-magnetic dipole ROA can be measured because the electric dipole-electric quadrupole component vanishes at this angle. Forward scattering,<sup>33</sup> scattering circular polarization<sup>34</sup> (SCP) and dual circular polarization<sup>35</sup> (DCP) modulation schemes were realized although they are not as favourable as the incident circular polarization (ICP) backscattering geometry for enhancing the signal to noise ratio. They do have merits in identifying ROA mechanism in forward scattering, and in reducing artifacts from strongly polarized Raman bands in DCP in backscattering in which the ROA intensity does not include the contribution from the isotropic scattering component which is the major source of ROA artifacts. Another ROA measurement strategy was proposed based on recent developments in FT Raman technology,<sup>36</sup> but it seems unlikely to compete with the multichannel CCD Raman spectrometer for instrument sensitivity at the moment although FT-Raman does have the advantage that it is free of fluorescence interference.

Parallel to the advance in instrumentation, a breakthrough has also taken place in ROA theory. By evaluating molecular polarizability and optical activity derivatives, reliable Raman and ROA intensities can be obtained on using quantum mechanical *ab initio* methods.<sup>37, 38</sup> A variety of small chiral molecules have been chosen to compare their *ab initio* results with experimental spectra. Prediction of ROA spectra from these calculation are in good agreement with data obtained from ROA experiment. This will lead eventually to the interpretation of ROA spectra with a firm theoretical base. Even more encouraging, *ab initio* calculations

of the amino acid L-alanine<sup>39</sup> have been carried out with results that compare favourably with experiment.

I came into the ROA field at the beginning of the new phase of ROA development and was involved in the development of the new ROA instrument and explored its application to biological molecules. This thesis is focused on the first ROA studies on a range of peptides, polypeptides, proteins and carbohydrates and will provide the basis for further ROA investigations of complex biological molecules.

## REFERENCES

1. S. F. Mason, *Molecular Optical Activity and Chiral Discrimination*. Cambridge University Press, Cambridge, 1982.
2. L. D. Barron, *Molecular Light Scattering and Optical Activity*. Cambridge University Press, Cambridge, 1982.
3. H. B. Kagan, *Stereochemistry and Fundamental and Methods*. Vol.2, Georg Thieme Publishers, Stuttgart, 1977.
4. C. K. Methews and K. E. van Holde, *Biochemistry*, Benjamin/Cummings Publishing Company, Inc, Readwood City, 1990.
5. S. Allenmark, *Chromatographic Enantioseparation methods and Application*, Second Edition. Ellis Horwood, New York, 1991.
6. N. Greenfield and G. D. Fasman, *Biochemistry*, **8**, 4108, 1969.
7. F. M. Pohl, *Nature*, **260**, 365, 1976.
8. J. W. Lewis, R. A. Goldbeck, D. S. Kliger, X. L. Xie, R. C. Dunn and J. D. Simon, *J. Phys. Chem.*, **96**, 5243, 1992.
9. R. Woody, *J. Polym. Sci., Macromol. Rev.*, **12**, 181, 1977.
10. E. C. Hsu and G. Holzwarth, *J. Chem. Phys.*, **59**, 4678, 1973.
11. G. Holzwarth, E. C. Hsu, H. S. Mosher, T. R. Faulkner and A. Moscowitz, *J. Am. Chem. Soc.*, **96**, 251, 1974.
12. L. A. Nafie, J. C. Cheng and P. J. Stephens, *J. Am. Chem. Soc.*, **97**, 3842, 1975.
13. T. A. Keiderling, *Appl. Spectrosc. Rev.*, **17**(2)., 189, 1981.
14. T. A. Keiderling, in *Practical Fourier Transform Infrared Spectroscopy*, edited by J. R. Farraro and K. Krishnan, Academic Press, New York,

203, 1990.

15. P. J. Stephens and M. A. Lowe, *Annul. Rev. Phys. Chem.*, **36**, 213, 1985.
16. P. J. Stephens, *J. Phys. Chem.*, **89**, 748, 1985
17. P. J. Stephens, *J. Phys. Chem.*, **91**, 1712, 1987.
18. P. Pancoska, S. C. Yasui and T. A. Keiderling, *Biochemistry*, **28**, 5917, 1989.
19. Y. C. Keiderling and T. A. Keiderling, *J. Am. Chem. Soc.*, **108**, 5576, 1986.
20. P. W. Atkins and L. D. Barron, *Mol. Phys.*, **16**, 453, 1969.
21. L. D. Barron and A. D. Buckingham, *Mol. Phys.*, **20**, 1111, 1971.
22. L. D. Barron, M. P. Bogaard and A. D. Buckingham, *J. Am. Chem. Soc.*, **95**, 603, 1973.
23. W. Hug, *Appl. Spectrosc.*, **35**, 115, 1981.
24. L. D. Barron and J. Vrbancich, *J. Raman Spectrosc.*, **15**, 47, 1984.
25. L. Hecht, B. Jordanov and B. Schrader, *Appl. Spectrosc.*, **41**, 295, 1987.
26. L. A. Nafie, in *Advances in Infrared and Raman spectroscopy*. edited by R. J. H. Clark and R. E. Hester, **11**, 123, 1984.
27. L. D. Barron and J. Vrbancich, *Top. Cur. Chem.*, **123**, 151, 1984.
28. L. Hecht, L. D. Barron and W. Hug, *Chem. Phys. Lett.*, **158**, 341, 1989.
29. L. D. Barron, L. Hecht, W. Hug and M. J. MacIntosh, *J. Am. Chem. Soc.*, **111**, 8731, 1989.
30. M. J. Pelletier and R. C. Reeder, *Appl. Spectrosc.*, **45**, 765, 1991.
31. L. Hecht, L. D. Barron, A. R. Gargaro, Z. Q. Wen and W. Hug, *J. Raman Spectrosc.*, **23**, 401, 1992.
32. L. Hecht and L. D. Barron, *Spectrochim. Acta*, **45A**, 671, 1989.

33. L. D. Barron, L. Hecht, A. R. Gargaro, and W. Hug, *J. Raman Spectrosc.*, **21**, 375, 1990.
34. K. M. Spencer, T. B. Freedman and L. A. Nafie, *Chem. Phys. Lett.*, **149**, 367, 1988.
35. L. A. Nafie and T. B. Freedman, *Chem. Phys. Lett.*, **154**, 260, 1989.
36. P. L. Polavarapu, *Spectrochim. Acta*, **46A**, 171, 1990.
37. R. D. Amos, *Chem. Phys. Lett.*, **87**, 23, 1982.
38. P. L. Polavarapu, *J. Phys. Chem.*, **94**, 8106, 1990.
39. L. D. Barron, A. R. Gargaro, L. Hecht and P. L. Polavarapu, *Spectrochim. Acta*, **47A**, 1001, 1991.

## CHAPTER 2

# Basic theory of Raman optical activity

### 2.1 Introduction

The theory of Raman optical activity will be briefly reviewed in this chapter. ROA theory can be roughly divided into two different levels. At the fundamental level, the ROA theory has been well established based on the semiclassical treatment of the interaction of light with matter.<sup>1</sup> The light radiation is considered as a classical electromagnetic wave whereas the molecules are treated as quantum mechanical objects. It revealed that the light waves scattered via the polarizability and optical activity tensors of a chiral molecule interfere with each other, thereby generating a small differential scattered intensity when the molecules are irradiated by right and left circularly polarized incident light.

A further development comprises the original ROA theory and the ROA dependence on the polarization and propagation vectors of the incident and scattered photons.<sup>2</sup> This allows an explicit description of ROA theory including the angular dependence and the different experimental geometries. An elegant analysis of all ROA measurement geometries based on this complete ROA theory reveals that backscattering is the best experimental strategy.<sup>3</sup> Another advance is that a new form of ROA, referred to as linear polarization ROA, has been proposed based on the

less restricted levels of approximation.<sup>4</sup>

At the more advanced level, ROA theory is expected to be able to interpret the ROA spectra of molecules and establish the correlation between ROA spectra and molecular structure and conformation. This is the ultimate goal of ROA theory development and is an even more challenging task. Nevertheless, the *ab initio* quantum-mechanical method has been developed to predict Raman intensity, normal vibrational mode frequency and ROA intensity and signs.<sup>5</sup> This opens up the possibility of interpreting ROA spectra at a fundamental level rather than relying upon empirical conclusions. A number of small chiral molecules have been tested by this *ab initio* quantum mechanical approach.<sup>6, 7</sup> Moreover, ROA spectra of a few simple model biological molecules such as L-alanine and tartaric acid have been calculated by the *ab initio* ROA theory and the results agree well with the experimental spectra.<sup>8, 9</sup>

Since ROA theory involves second order perturbation in terms of quantum mechanics and the expressions are extremely complex, Cartesian tensor notation is employed to treat the interaction between electromagnetic radiation and the molecular optical property tensors of the molecule.

## **2.2 Molecular multipole moments**

It is appropriate to begin with the concept of molecular multipole moments of a molecule, which can be considered as an object composed of charges and masses, in order to deal with light scattering by molecules.

The physical nature of the light scattering is the emission of a characteristic second light wave from the oscillating electric and magnetic dipole moments of the molecule induced by the electromagnetic fields of the incident light wave. Three basic electric and magnetic multipole moments are encountered in the Raman optical activity theory and are defined as follows:

$$\mu_{\alpha} = \sum_i e_i r_{i\alpha} \quad 2.2.1$$

$$m_{\alpha} = \sum_i \frac{e_i}{m_i} \varepsilon_{\alpha\beta\gamma} r_{i\beta} p_{i\gamma} \quad 2.2.2$$

$$\Theta_{\alpha\beta} = \frac{1}{2} \sum_i e_i (3r_{i\alpha} r_{i\beta} - r_i^2 \delta_{\alpha\beta}) \quad 2.2.3$$

where  $\mu_{\alpha}$  is the electric dipole moment or the first moment of a collection of charges,  $m_{\alpha}$  is the magnetic dipole moment, which is the first moment of a circulating current distribution and  $\Theta_{\alpha\beta}$  is the electric quadrupole moment or the second moment of a collection of charges;  $e_i$ ,  $m_i$ ,  $p_{i\alpha}$  are the electric charge, the corresponding mass and the linear momentum at a distance  $r_{i\alpha}$  respectively;  $\delta_{\alpha\beta}$  is the unit second rank symmetric tensor or Kronecka delta symbol, and  $\varepsilon_{\alpha\beta\gamma}$  is the unit third rank antisymmetric tensor, which has a value 1 or -1 if  $\alpha\beta\gamma$  is an even or odd permutation of the order of XYZ and zero for any other combination.

### 2.3 Light scattering by a molecule

The incident light that interacts with the molecule can be generally

treated as electromagnetic radiation with an angular frequency  $\omega$  propagating in the direction of a unit vector  $\mathbf{n}$  and expressed as

$$\tilde{\mathbf{E}}_{\alpha} = \tilde{\mathbf{E}}_{\alpha}^{(\omega)} \exp[-i\omega(t - \mathbf{n}_{\beta} \mathbf{r}_{\beta} / c)] \quad 2.3.1a$$

$$\tilde{\mathbf{B}}_{\alpha} = \frac{1}{c} \epsilon_{\alpha\beta\gamma} n_{\beta} \tilde{\mathbf{E}}_{\gamma} \quad 2.3.1b$$

where  $\tilde{\mathbf{E}}_{\alpha}$  is the complex electric vector of the light wave and  $\tilde{\mathbf{B}}_{\alpha}$  is the magnetic counterpart (a tilde is used to denote a complex quantity), and  $c$  is the velocity of light. Fig 2.1 displays a schematic view of the scattering geometry of a molecule in a radiation field.

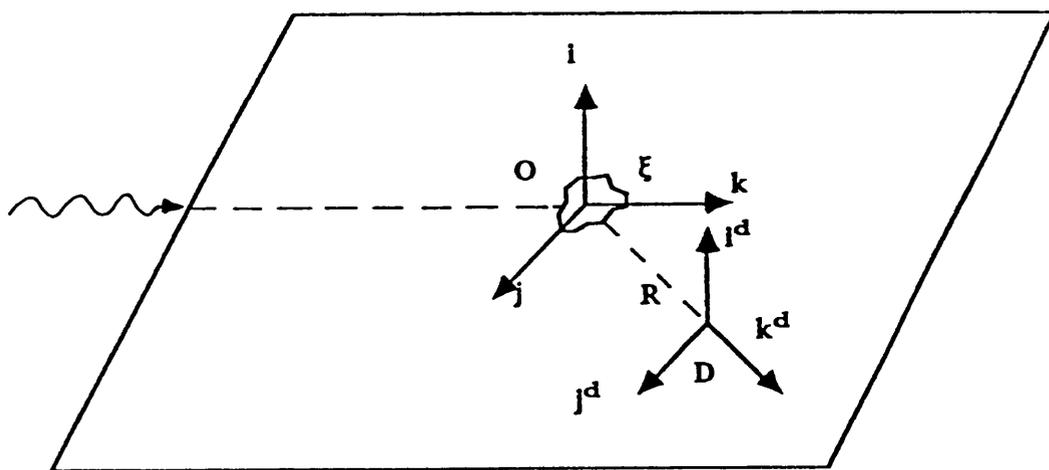


Fig. 2.1 A molecule in a radiation field. O is the molecular origin and R is the distance between the molecule and the detector D.

The scattered light by a molecule received by a detector at a position  $\mathbf{d}$ , a distance  $R$  from the molecular origin is, according to classical electrodynamics,<sup>1</sup>

$$\begin{aligned} \tilde{\mathbf{E}}_{\alpha}^{\mathbf{d}} = & \frac{\omega\mu}{4\pi R} \exp[iR/c-t] [\mu_{\alpha}^{\sim(\circ)} - n_{\alpha}^{\mathbf{d}} n_{\beta}^{\mathbf{d}} \mu_{\beta}^{\sim(\circ)} - \frac{1}{c} \epsilon_{\alpha\beta\gamma} n_{\beta}^{\mathbf{d}} m_{\gamma}^{\sim(\circ)} \\ & - \frac{i\omega}{3c} (n_{\beta}^{\mathbf{d}} \tilde{\Theta}_{\alpha\beta}^{\sim(\circ)} - n_{\alpha}^{\mathbf{d}} n_{\beta}^{\mathbf{d}} n_{\gamma}^{\mathbf{d}} \tilde{\Theta}_{\beta\gamma}^{\sim(\circ)}) + \dots ] \end{aligned} \quad 2.3.2$$

where  $n_{\alpha}^{\mathbf{d}}$  is the propagation vector of the detected wave.

For a molecule in a radiation field, to characterize the response of a molecule to the electromagnetic field of the light wave, dynamic molecular property tensors should be invoked. The induced multipole moments are expressed by<sup>1</sup>

$$\tilde{\mu}_{\alpha}^{\sim(\circ)} = ( \alpha_{\alpha\beta}^{\sim} + \frac{1}{c} \epsilon_{\gamma\delta\beta} n_{\delta}^{\mathbf{i}} \tilde{G}_{\alpha\gamma}^{\sim} + \frac{i\omega}{3c} n_{\gamma}^{\mathbf{i}} \tilde{A}_{\alpha\gamma\beta}^{\sim} + \dots ) \tilde{E}_{\beta}^{\sim(\circ)} \quad 2.3.3a$$

$$m_{\alpha}^{\sim(\circ)} = g_{\alpha\beta}^{\sim} \tilde{E}_{\beta}^{\sim(\circ)} + \dots \quad 2.3.3b$$

$$\tilde{\Theta}_{\alpha\beta}^{\sim(\circ)} = \tilde{A}_{\gamma\alpha\beta}^{\sim} \tilde{E}_{\gamma}^{\sim(\circ)} + \dots \quad 2.3.3c$$

where  $n_{\alpha}^{\mathbf{i}}$  is the propagation vector of the incident wave. The scattered wave at the detector position is now

$$\tilde{\mathbf{E}}^{\mathbf{d}} = \frac{\omega^2 \mu_0}{4\pi R} \exp[i\omega(R/c-t)] \tilde{a}_{\alpha\beta}^{\sim} \tilde{\mathbf{E}}_{\beta}^{\sim(\circ)} \quad 2.3.4$$

and  $\tilde{a}_{\alpha\beta}$  is a scattering tensor:

$$\tilde{a}_{\alpha\beta} = \tilde{\alpha}_{\alpha\beta} + \frac{i\omega}{3c} (n_{\gamma}^i \tilde{A}_{\alpha\gamma\beta} - n_{\gamma}^d \tilde{A}_{\beta\gamma\alpha}) + \frac{1}{c} (\epsilon_{\gamma\delta\beta} n_{\delta}^i \tilde{G}_{\alpha\gamma} + \epsilon_{\gamma\delta\alpha} n_{\delta}^d \tilde{g}_{\gamma\beta}) + \dots \quad 2.3.5$$

where  $\alpha_{\alpha\beta}$ ,  $G'_{\alpha\beta}$ ,  $A_{\alpha\beta\gamma}$  are dynamic molecular property tensors obtained from time dependent perturbation theory<sup>10</sup> given by

$$\alpha_{\alpha\beta} = \frac{2}{\hbar} \sum_{j \neq n} \left( \frac{\omega_{jn}}{\omega_{jn} - \omega^2} \right) \text{Re}(\langle n | \mu_{\alpha} | j \rangle \langle j | \mu_{\beta} | n \rangle) \quad 2.3.6a$$

$$G'_{\alpha\beta} = - \frac{2}{\hbar} \sum_{j \neq n} \left( \frac{\omega_{jn}}{\omega_{jn} - \omega^2} \right) \text{Im}(\langle n | \mu_{\alpha} | j \rangle \langle j | m_{\beta} | n \rangle) \quad 2.3.6b$$

$$A_{\alpha\beta\gamma} = \frac{2}{\hbar} \sum_{j \neq n} \left( \frac{\omega_{jn}}{\omega_{jn} - \omega^2} \right) \text{Re}(\langle n | \mu_{\alpha} | j \rangle \langle j | \Theta_{\beta\gamma} | n \rangle) \quad 2.3.6c$$

$\omega_{jn} = \omega_j - \omega_n$  is the angular frequency between the ground state  $|n\rangle$  and the  $j$ th excited state  $|j\rangle$ .

The electric dipole-electric dipole tensor  $\alpha_{\alpha\beta}$  is the ordinary symmetric polarizability that contributes to light scattering and refraction phenomena. All molecules support this polarizability.  $G'_{\alpha\beta}$  is the electric dipole-magnetic dipole optical activity tensor, the isotropic part of which is responsible for natural optical rotation of isotopic samples.  $A_{\alpha\beta\gamma}$  is the electric dipole-electric quadrupole optical activity tensor, which makes additional contribution to optical rotation in oriented samples. These polarizability and optical activity tensors are valid when the incident light is far from any absorption frequency  $\omega_{jn}$  of the molecules.

## 2.4 Basic Raman optical activity equation

The Raman optical activity is measured by a small difference in the scattered Raman intensity under right and left circularly polarized incident light. A useful experimental quantity is the dimensionless Raman circular intensity difference (CID), which is defined as<sup>11</sup>

$$\Delta_{\alpha} = (I_{\alpha}^R - I_{\alpha}^L) / (I_{\alpha}^R + I_{\alpha}^L) \quad 2.4.1$$

where  $\alpha$  denotes polarization of the scattered intensity, and  $I^R$  and  $I^L$  are the scattered intensities in right and left circularly polarized incident light.

The following CID equations are obtained by developing an expression from classical electrodynamics for the electric field vector radiated by the oscillating electric dipole, magnetic dipole and electric quadrupole moments induced by circularly polarized light waves and then calculating the corresponding intensity:<sup>1</sup>

$$\Delta(0^{\circ}) = \frac{4(3\alpha_{\alpha\alpha}G'_{\beta\beta} + \alpha_{\alpha\beta}G'_{\alpha\beta} - \frac{1}{3}\omega\alpha_{\alpha\beta}\epsilon_{\alpha\gamma\delta}A_{\gamma\delta\beta})}{c(7\alpha_{\lambda\mu}\alpha_{\lambda\mu} + \alpha_{\lambda\lambda}\alpha_{\mu\mu})} \quad 2.4.2a$$

$$\Delta(180^{\circ}) = \frac{8(3\alpha_{\alpha\beta}G'_{\alpha\beta} - \alpha_{\alpha\alpha}G'_{\beta\beta} + \frac{1}{3}\omega\alpha_{\alpha\beta}\epsilon_{\alpha\gamma\delta}A_{\gamma\delta\beta})}{c(7\alpha_{\lambda\mu}\alpha_{\lambda\mu} + \alpha_{\lambda\lambda}\alpha_{\mu\mu})} \quad 2.4.2b$$

$$\Delta_x(90^{\circ}) = \frac{2(7\alpha_{\alpha\beta}G'_{\alpha\beta} + \alpha_{\alpha\alpha}G'_{\beta\beta} + \frac{1}{3}\omega\alpha_{\alpha\beta}\epsilon_{\alpha\gamma\delta}A_{\gamma\delta\beta})}{c(7\alpha_{\lambda\mu}\alpha_{\lambda\mu} + \alpha_{\lambda\lambda}\alpha_{\mu\mu})} \quad 2.4.2c$$

$$\Delta_z(90^\circ) = \frac{4(3\alpha_{\alpha\beta}G'_{\alpha\beta} - \alpha_{\alpha\alpha}G'_{\beta\beta} - \frac{1}{3}\omega\alpha_{\alpha\beta}\varepsilon_{\alpha\gamma\delta}A_{\gamma\delta\beta})}{2c(3\alpha_{\lambda\mu}\alpha_{\lambda\mu} - \alpha_{\lambda\lambda}\alpha_{\mu\mu})} \quad 2.4.2d$$

It is beneficial to rewrite these CID equations in terms of the following tensor invariants with respect to molecule fixed axes X, Y, Z:

$$\alpha = \frac{1}{3} \alpha_{\alpha\alpha} = \frac{1}{3} (\alpha_{xx} + \alpha_{yy} + \alpha_{zz}) \quad 2.4.3a$$

$$\beta(\alpha)^2 = \frac{1}{2} (3\alpha_{\alpha\beta}\alpha_{\alpha\beta} - \alpha_{\alpha\alpha}\alpha_{\beta\beta}) \quad 2.4.3b$$

$$G' = \frac{1}{3} G'_{\alpha\alpha} = (G'_{xx} + G'_{yy} + G'_{zz}) \quad 2.4.3c$$

$$\beta(G')^2 = \frac{1}{2} (3\alpha_{\alpha\beta}G'_{\alpha\alpha} - \alpha_{\alpha\alpha}G'_{\beta\beta}) \quad 2.4.3d$$

$$\beta(A)^2 = \frac{1}{2} \omega\alpha_{\alpha\beta}\varepsilon_{\alpha\gamma\delta}A_{\gamma\delta\beta} \quad 2.4.3e$$

The CIDs now become:

$$\Delta(0^\circ) = \frac{8[45\alpha G' + \beta(G')^2 - \beta(A)^2]}{2c[45\alpha^2 + 7\beta(\alpha)^2]} \quad 2.4.4a$$

$$\Delta(180^\circ) = \frac{48[\beta(G')^2 + 1/3\beta(A)^2]}{2c[45\alpha^2 + 7\beta(\alpha)^2]} \quad 2.4.4b$$

$$\Delta_x(90^\circ) = \frac{2[45\alpha G' + 7\beta(G')^2 + \beta(A)^2]}{c[45\alpha^2 + 7\beta(\alpha)^2]} \quad 2.4.4c$$

$$\Delta_z(90^\circ) = \frac{12[\beta(G)^2 - 1/3\beta(A)^2]}{6c\beta(\alpha)^2} \quad 2.4.4d$$

These expressions apply explicitly to Rayleigh optical activity. For Raman optical activity, the same expressions can be used but the molecular property tensors must be replaced by corresponding vibrational Raman transition tensors that take into account of the initial and final vibrational states of the molecules.

## 2.5 Molecular scattering transition tensors

The complex transition polarizability and optical activity tensors are derived from time-dependent perturbation theory<sup>12</sup> and expressed as:

$$(\tilde{\alpha}_{\alpha\beta})_{mn} = \frac{1}{\hbar} \sum_{j \neq n, m} \left[ \frac{\langle m | \mu_\alpha | j \rangle \langle j | \mu_\beta | n \rangle}{\omega_{jn} - \omega} + \frac{\langle m | \mu_\beta | j \rangle \langle j | \mu_\alpha | n \rangle}{\omega_{jn} + \omega} \right] \quad 2.5.1a$$

$$(\tilde{G}_{\alpha\beta})_{mn} = \frac{1}{\hbar} \sum_{j \neq n, m} \left[ \frac{\langle m | \mu_\alpha | j \rangle \langle j | m_\beta | n \rangle}{\omega_{jn} - \omega} + \frac{\langle m | m_\beta | j \rangle \langle j | \mu_\alpha | n \rangle}{\omega_{jn} + \omega} \right] \quad 2.5.1b$$

$$(\tilde{A}_{\alpha\beta\gamma})_{mn} = \frac{1}{\hbar} \sum_{j \neq n, m} \left[ \frac{\langle m | \mu_\alpha | j \rangle \langle j | \theta_{\beta\gamma} | n \rangle}{\omega_{jn} - \omega} + \frac{\langle m | \theta_{\beta\gamma} | j \rangle \langle j | \mu_\alpha | n \rangle}{\omega_{jn} + \omega} \right] \quad 2.5.1c$$

$$(\tilde{g}_{\alpha\beta})_{mn} = \frac{1}{\hbar} \sum_{j \neq n, m} \left[ \frac{\langle m | m_\alpha | j \rangle \langle j | \mu_\beta | n \rangle}{\omega_{jn} - \omega} + \frac{\langle m | \mu_\beta | j \rangle \langle j | m_\alpha | n \rangle}{\omega_{jn} + \omega} \right] \quad 2.5.1d$$

$$(\tilde{A}_{\gamma\alpha\beta})_{mn} = \frac{1}{\hbar} \sum_{j \neq n, m} \left[ \frac{\langle m | \Theta_{\alpha\beta} | j \rangle \langle j | \mu_{\gamma} | n \rangle}{\omega_{jn} - \omega} + \frac{\langle m | \mu_{\gamma} | j \rangle \langle j | \Theta_{\alpha\beta} | n \rangle}{\omega_{jn} + \omega} \right] \quad 2.5.1e$$

At transparent frequencies, the required vibrational Raman transition tensors can be obtained within Placzek's polarizability theory which considers the dependence of the ground state electronic polarizability on the normal coordinates of vibration.<sup>13</sup> Only a fundamental vibrational transition is considered and the molecule is in its electronic ground state and the rotational contributions are ignored with the assumption that the initial and final vibrational states  $|n_v\rangle$  and  $|m_v\rangle$  are real. The vibrational transition tensors can be written:

$$(\tilde{\alpha}_{\alpha\beta})_{mn} \longrightarrow \langle m_v | \alpha_{\alpha\beta}(Q) | n_v \rangle \quad 2.5.2a$$

$$(\tilde{G}_{\alpha\beta})_{mn} \longrightarrow \langle m_v | G'_{\alpha\beta}(Q) | n_v \rangle \quad 2.5.2b$$

$$(\tilde{A}_{\alpha\beta\gamma})_{mn} \longrightarrow \langle m_v | A_{\alpha\beta\gamma}(Q) | n_v \rangle \quad 2.5.2c$$

$$(\tilde{g}_{\alpha\beta})_{mn} \longrightarrow \langle m_v | G'_{\beta\gamma}(Q) | n_v \rangle \quad 2.5.2d$$

$$(A_{\gamma\alpha\beta})_{mn} \longrightarrow \langle m_v | A_{\gamma\alpha\beta}(Q) | n_v \rangle \quad 2.5.2e$$

where  $\alpha_{\alpha\beta}(Q)$  etc. are interpreted as effective operators which depend parametrically on the normal vibrational coordinates  $Q$  and so bring about the vibrational transition. By expanding  $\alpha_{\alpha\beta}(Q)$  etc. in a Taylor series in  $Q$ , the Raman intensity and optical activity in a fundamental

vibrational transition  $1_p \leftarrow 0$  associated with the normal vibrational coordinate  $Q_p$  are found to be:<sup>1</sup>

$$\langle 0 | \alpha_{\alpha\beta} | 1_p \rangle \langle 1_p | \alpha_{\alpha\beta} | 0 \rangle = \left( \frac{\hbar}{2\omega_p} \right) \left( \frac{\partial \alpha_{\alpha\beta}}{\partial Q_p} \right)_o \left( \frac{\partial \alpha_{\alpha\beta}}{\partial Q_p} \right)_o \quad 2.5.3a$$

$$\langle 0 | \alpha_{\alpha\beta} | 1_p \rangle \langle 1_p | G'_{\alpha\beta} | 0 \rangle = \left( \frac{\hbar}{2\omega_p} \right) \left( \frac{\partial \alpha_{\alpha\beta}}{\partial Q_p} \right)_o \left( \frac{\partial G'_{\alpha\beta}}{\partial Q_p} \right)_o \quad 2.5.3b$$

$$\langle 0 | \alpha_{\alpha\beta} | 1_p \rangle \langle 1_p | \epsilon_{\alpha\gamma\beta} A_{\gamma\delta\beta} | 0 \rangle = \left( \frac{\hbar}{2\omega_p} \right) \left( \frac{\partial \alpha_{\alpha\beta}}{\partial Q_p} \right)_o \epsilon_{\alpha\gamma\beta} \left( \frac{\partial A_{\gamma\delta\beta}}{\partial Q_p} \right)_o \quad 2.5.3c$$

## 2.6 The bond polarizability model of ROA and the merits of backscattering

The bond polarizability model of ROA theory is based on Placzek's approximation<sup>13</sup> and the Raman intensity theory developed by Wolkenstein.<sup>14</sup> The central idea is to treat the total molecular polarizability of a molecule as the tensor sum of independent bonds. The bond's polarizability tensor only depends upon the internal coordinate of the bonds and is independent of other components. In other words, the bond polarizability model break down the molecule into pieces that are either its constituent atoms or its constituent bonds, respectively.

By taking account of the origin-dependence of  $G'_{\alpha\beta}$  and  $A_{\alpha\beta\gamma}$ , the expressions for the molecular polarizability and optical activity tensors within the bond polarizability theory are found to be:<sup>1</sup>

$$\alpha_{\alpha\beta} = \sum_i \alpha_{i,\alpha\beta} \quad 2.6.1a$$

$$G'_{\alpha\beta} = \sum_i (G'_{i\alpha\beta} - 1/2 \omega \epsilon_{\beta\gamma\delta} R_{i\gamma} \alpha_{i\alpha\gamma}) \quad 2.6.1b$$

$$A_{\alpha\beta\gamma} = \sum_i (A_{i\alpha\beta\gamma} + 3/2 R_{i\gamma} \alpha_{i\beta\alpha} + 3/2 R_{i\beta} \alpha_{i\gamma\alpha} - R_{i\delta} \alpha_{\delta\alpha} \delta_{\beta\gamma}) \quad 2.6.1c$$

where  $R_{i\alpha}$  is the vector from the molecular origin to the  $i$ th bond or group,  $\alpha_{i\alpha\beta}$ ,  $G'_{i\alpha\beta}$  and  $A_{i\alpha\beta\gamma}$  are the corresponding tensors of the individual bonds.

The model has proved to be useful, in particular, in providing simple physical pictures of the source of ROA in typical chiral structure. A further assumption in the bond polarizability model was made that all bonds have axial symmetry, in which case contributions to the total polarizability and optical activity tensors from intrinsic components vanish, and

$$\alpha_{\alpha\beta} G'_{\alpha\beta} = 1/3 \omega \alpha_{\alpha\beta} \epsilon_{\alpha\gamma\delta} A_{\gamma\delta\beta} \quad 2.6.2a$$

$$\alpha_{\alpha\alpha} G'_{\beta\beta} = 0 \quad 2.6.2b$$

The general ROA CIDs equations (2.4) now take the following form (since  $\beta(G')^2 = \beta(A)^2$  and  $\alpha G' = 0$ ):

$$\Delta(0^\circ) = 0 \quad 2.6.3a$$

$$\Delta(180^\circ) = \frac{64\beta(G')^2}{c[45\alpha^2 + 7\beta(\alpha)^2]} \quad 2.6.3b$$

$$\Delta_x(90^\circ) = \frac{16\beta(G')^2}{c[45\alpha^2 + 7\beta(\alpha)^2]} \quad 2.6.3c$$

$$\Delta_z(90^\circ) = \frac{8\beta(G)^2}{6c\beta(\alpha)^2} \quad 2.6.3d$$

One important conclusion that can be immediately inferred from a comparison between equation 2.6.3b and 2.6.3c is that the ROA intensity in backscattering is four times larger than that in polarized right angle scattering with the corresponding conventional Raman intensity increased two fold, which means a  $2\sqrt{2}$  fold signal to noise ratio (SNR) enhancement for the ROA measurements at the same recording time.<sup>16</sup> In other words, for a given SNR, the measurement speed is increased eight times faster under the same measurement condition (i.e. same laser power, collection efficiency, quantum efficiency of the detector and so on).

This prediction has been confirmed experimentally for pure saturated hydrocarbons like  $\alpha$ -trans-pinane.<sup>16</sup> Another merit of backscattering geometry is the suppression of the artifacts which had plagued ROA measurement in polarized right angle scattering. Because the backscattering geometry creates an almost perfect rotational symmetry with respect to the scattering and collection cone, thus the possible artifacts caused by the imperfection of the circular polarization of the probing beam are minimized. Analysis of all possible ROA measurement geometries, together with instrumentation feasibility considerations, have identified the backscattering as the optimum experimental strategy.<sup>3</sup>

## 2.7 The *ab initio* ROA calculation

The ultimate goal of ROA theory development is the prediction of the

ROA spectrum from a known molecular structure, or even better to predict the structure from the measured ROA spectrum. Simple models were proposed to interpret ROA spectra as initial attempts in the early days. The bond polarizability model<sup>17</sup> was one of them. Other simple model includes the atomic dipole interaction model<sup>18</sup> and the orbital polarization model<sup>19</sup> based on molecular orbital calculations. Unfortunately, none of them can give a successful prediction of ROA spectra as these models require parametrization of the bond or atom polarizability although they have given some insight into the mechanism of generating ROA from typical chiral structure.

The most promising ROA theory at present appears to be the *ab initio* calculation realized by Polavarapu<sup>5</sup> based on the recent advance made by Amos<sup>20, 21</sup> for evaluating the optical activity tensor  $G'_{\alpha\beta}$  and  $A_{\alpha\beta\gamma}$ . The theoretical basis of *ab initio* ROA calculations is that, as given in equations (2.5), the Raman and ROA intensity for a given vibrational normal mode arise from the products of polarizability and optical activity tensor derivatives with respect to normal mode coordinates. The required tensor derivatives  $(\partial\alpha_{\alpha\beta}/\partial Q_p)_o$  and  $(\partial A_{\alpha\beta\gamma}/\partial Q_p)_o$  can now be calculated analytically in the *ab initio* CADPAC program implemented by Amos<sup>22</sup>, although  $(\partial G'_{\alpha\beta}/\partial Q_p)_o$  has to be calculated numerically.

The quantum mechanical expression for the polarizability and optical activity were given in equations 2.3.6. A further simplification is the replacement of  $\omega_{jn}^2 - \omega^2$  by  $\omega^2$  in the denominators, which is a reasonable approximation for Raman scattering at transparent wavelengths ( $\omega_{jn} \gg \omega$ ):

$$\alpha_{\alpha\beta} = 2 \sum_{j \neq n} \frac{\langle n | \mu_{\alpha} | j \rangle \langle j | \mu_{\beta} | n \rangle}{E_j - E_n} \quad 2.7.1a$$

$$A_{\alpha\beta\gamma} = 2 \sum_{j \neq n} \frac{\langle n | \mu_{\alpha} | j \rangle \langle j | \Theta_{\beta\gamma} | n \rangle}{E_j - E_n} \quad 2.7.1b$$

These quantities can be evaluated by using the first order correction  $\psi_n^{(1)}(E_{\beta})$  to the electronic wavefunction in the presence of a static electric field. They can be expressed by

$$\alpha_{\alpha\beta} = 2 \langle \psi_n^{(0)} | \mu_{\alpha} | \psi_n^{(1)}(E_{\beta}) \rangle \quad 2.7.2a$$

$$A_{\alpha\beta\gamma} = 2 \langle \psi_n^{(0)} | \Theta_{\beta\gamma} | \psi_n^{(1)}(E_{\beta}) \rangle \quad 2.7.2b$$

The electric dipole-magnetic dipole optical activity tensor can not be calculated analytically because it vanishes when  $\omega \rightarrow 0$  and so does not have a static limit. However,  $1/\omega G'_{\alpha\beta}$  does have a static limit which can be written in the form

$$(1/\omega G'_{\alpha\beta})_{\omega=0} = -2\hbar \text{Im}(\langle \psi_n^{(1)}(E_{\alpha}) | \psi_n^{(1)}(B_{\beta}) \rangle) \quad 2.7.3$$

where  $\psi_n^{(1)}(B_{\beta})$  refers to the corresponding wavefunction perturbed by a static magnetic field. By expressing the wavefunctions in terms of molecular orbitals, it is possible to calculate the polarizability and optical activity tensors from two coupled Hartree-Fock SCF calculations, one to give the molecular orbitals perturbed by a static electric field, the other

by a static magnetic field. The required tensor derivatives are then calculated by using the above results to evaluate  $\alpha_{\alpha\beta}$ ,  $(1/\omega)G'_{\alpha\beta}$  and  $A_{\alpha\beta\gamma}$  in their static limits at the equilibrium geometry and at the geometries displaced by  $0.005\text{\AA}$  along each atomic coordinate with the CADPAC program.

## REFERENCES

1. L. D. Barron, *Molecular Light Scattering and Optical Activity*, Cambridge University Press, Cambridge, 1982.
2. L. Hecht and L. A. Nafie, *Mol. Phys.*, **72**, 441, 1991.
3. L. Hecht and L. D. Barron, *Appl. Spectrosc.*, **44**, 483, 1990.
4. L. A. Nafie and L. Hecht, *Chem. Phys. Lett.*, **174**, 575, 1991.
5. P. L. Polavarapu, *J. Phys. Chem.*, **94**, 8106, 1990.
6. P. K. Bose and P. L. Polavarapu, *Chem. Phys. Lett.*, **152**, 39, 1988.
7. T. M. Black, P. K. Bose, P. L. Polavarapu, L. D. Barron and L. Hecht, *J. Am. Chem. Soc.*, **112**, 1479, 1990.
8. L. D. Barron, A. R. Gargaro, L. Hecht and P. L. Polavarapu, *Spectrochim. Acta*, **47A**, 1001, 1991.
9. L. D. Barron, A. R. Gargaro, L. Hecht, P. L. Polavarapu and H. Sugeta, *Spectrochimica Acta*, **48A**, 1051, 1992.
10. A. D. Buckingham, *Adv. Chem. Phys.*, **12**, 107, 1967.
11. L. D. Barron and A. D. Buckingham, *Mol. Phys.*, **20**, 1111, 1971.
12. L. D. Barron and J. Escribano, *Chem. Phys. Lett.*, **98**, 437, 1985.
13. G. Placzek, in 'Handbuch der Radiologie', edited by E. Marx, Leipzig, Akademische Verlagsgesellschaft, Vol. 5, Part2, 205, 1934.
14. M. W. Wolkenstein, *Comp. Rend. Acad. Sci., U.R.S.S.* **32**, 185, 1941.
15. L. D. Barron and B. P. Clark, *Mol. Phys.*, **46**, 839, 1982.
16. L. Hecht, L. D. Barron and W. Hug, *Chem. Phys. Lett.*, **158**, 341, 1989.
17. L. D. Barron, A. D. Buckingham, *J. Am. Chem. Soc.*, **96**, 4769, 1979.
18. P. L. Polavarapu and L. A. Nafie, *J. Chem. Phys.*, **70**, 5582, 1979.

19. P. L. Polavarapu, *J. Chem. Phys.*, **77**, 2273, 1982.
20. R. D. Amos, *Chem. Phys. Lett.*, **87**, 23, 1982.
21. E. B. Graham and R. E. Raab, *Mol. Phys.*, **52**, 185, 1984.
22. A. D. Amos and J. E. Rice, *CADPAC (The Cambridge Analytical Derivative Package)*. Issue, 4.0, Cambridge, 1987.

## CHAPTER 3

# The Raman optical activity instrument

### 3.1. Introduction

The instrument required for ROA measurement is, in principle, simple and composed of two parts: a conventional Raman spectrometer plus a light modulator to generate the required circularly polarized light. However, in practice, it faces a number of challenges because the ROA signal is extremely weak. The difficulty of measuring ROA spectra may be illustrated by an old Chinese proverb 'seeking for a needle in the sea'. There has not yet been any commercial ROA spectrometer available so far although most of the components constituting an ROA spectrometer can be purchased from the market. The ROA instrumentation is in a state of rapid flux. All the ROA instruments used for ROA study therefore have to be constructed in the laboratory by the researcher themselves.

The development of ROA instruments can be roughly divided into three phases in terms of the adoption of different light detectors, which is the crucial component determining Raman instrument sensitivity. The first generation of ROA instruments was based on a scanning Raman spectrometer with a photomultiplier tube as detector together with an electro-optical light modulator.<sup>1</sup> It generated a large number of ROA data of pure organic chiral molecules in spite of its low instrument SNR and

long acquisition time.

The employment of the optical multichannel array detector in Raman spectrometer represented the second generation of ROA instrumentation.<sup>2</sup> It measures simultaneously a segment of a Raman spectrum instead of recording a few wavenumbers of the spectrum at a time in the old scanning system. These new detectors increased by over two orders of magnitude the measurement speed of an ROA spectrum. Good quality ROA spectra in right angle scattering geometry could be reliably recorded in a time scale of a few hours instead of days. The ROA instruments developed by Hug et al.<sup>3</sup> and Barron et al.<sup>4</sup> were two good examples of that generation. Unfortunately, the instrument sensitivity was not high enough for the study of biological molecules in aqueous solution, so they were restricted to the study of small model compounds, usually pure organic liquids.

The adoption of a backscattering geometry<sup>5</sup> and the advent of CCD light detectors<sup>6</sup> promoted ROA instrument sensitivity to above the threshold level that renders ROA studies on biological molecules in aqueous solution possible. Other forms of measurement strategy have certain merits: the reduction of artifacts from strongly polarized Raman band in SCP ROA<sup>8</sup> and better theoretical SNR in DCP ROA.<sup>9</sup> With the goal of the application of ROA spectroscopy to biological molecules in mind, a complete new ROA spectrometer using the ICP (Incident Circular Polarization) backscattering strategy and upgrading almost every optical component in the optical train has been constructed in our laboratory. This new instrument GUROAS1 ( Glasgow University Raman Optical Activity Spectrometer 1 ) represents the third generation of ROA

spectrometer and will be described in detail in the following sections.<sup>10</sup>

### **3.2 Some basic considerations for ROA instrumentation**

ROA is a polarized spectroscopic method.<sup>11</sup> A right and left circularly polarized light source is a prerequisite, which is produced by a light modulator composed of a non-linear optical crystal KD\*P driven by a high voltage differential amplifier. The Raman spectrometer disperses respectively the scattered Raman signal from samples under the right and left circularly polarized light. A difference and a sum spectrum between these two polarized states are generated. The difference spectrum is the ROA and the sum spectrum is the parent Raman.

The extremely weak ROA signal demands that the spectrometer should have very high sensitivity which in turn requires high throughput of the spectrograph, highly efficient collection optics, high power of the incident light and high quantum efficiency of the light detector. To improve instrument sensitivity is above all the priority in ROA instrumentation.

Another barrier in ROA measurements is the artifacts, which is false scattering differences for right and left circularly polarized light due to the experimental set up and not to the genuine chiroptical property of the molecules.<sup>12</sup> It is particularly associated with the polarization property of the Raman bands. The strongest polarized Raman bands usually suffer the largest artifacts. The first successful ROA measurement was not achieved until the understanding of the major mechanism of the artifact. Many efforts have been made to eliminate the artifacts. The mechanisms

of generating ROA artifacts have now been better understood through qualitative<sup>12,13</sup> and quantitative<sup>14</sup> analysis of the entire ROA optical system. The principle source of artifacts derives from a deviation of the polarization of the probing light from precisely right and left circular. In particular, birefringence of the cell walls and of the electro-optical crystal of the modulator and the instability of the offset voltage applied to the modulator can introduce linearly polarized contaminants.<sup>14</sup> ROA spectra free of artifacts can now be obtained from most samples without very strongly polarized Raman band. For molecules having strongly polarized Raman bands, it is possible to control artifacts to below the noise level.

Stray light from Rayleigh scattering is the third obstacle in ROA measurements, which require much higher rejection of stray light than conventional Raman due to its extremely weak signal. Otherwise, the ROA spectrum would be overwhelmed by the much stronger Rayleigh stray light ( Rayleigh scattering is  $10^6$  times stronger than ROA effect). It is even more severe in a multichannel backscattering Raman spectrometer than in a single scanning system with right angle scattering geometry. Stray light becomes more problematic for biological molecules in aqueous solution than organic neat liquid and increases with the molecular size for biological macromolecules such as proteins and DNA and sample concentration. Double or even triple monochromator had been previously employed to reject stray light in order to measure low frequency ROA spectra.<sup>4</sup> An obvious disadvantage in a double or triple monochromator spectrograph is the low throughput of the Raman signal, hence reducing the instrument sensitivity.

The last, but not least, problem in ROA experiment is fluorescence interference from the sample. This is particularly severe in biological molecules where fluorescence originate from chromophores within the sample itself as well as from impurities. The former is unavoidable and the latter can be reduced by conventional physico-chemical purification methods. Long time exposure of the sample in the laser beam is the last resort to burn off the fluorescence. No sophisticated procedure yet can ensure complete elimination of fluorescence. The development of FT-Raman technology<sup>15</sup> has now been proved to be free of fluorescence in Raman measurement by shifting the exciting light from visible to infrared. However, the loss of Raman intensity due to the Rayleigh scattering law (light scattering intensity is proportional to  $1/\lambda^4$ ) is intolerable for ROA measurement although for conventional Raman measurement it has become very popular.

### 3.3 The optical configuration of the ROA instrument

The optical configuration of the ROA instrument is depicted in Fig. 3.1. The light beam emitted from a laser (L) transmits through a modulator (M) to become circularly polarized light. A focusing lens (F) focuses the light beam into the sample (S). The scattered Raman light is depolarized by a Lyot depolarizer and collimated by a triplet lens and oriented by a plane reflection mirror in the backscattering block (B) to the collection camera lens (C) which in turn directs the scattered Raman light into the entrance slit of the spectrograph. An optical edge filter (F) is set in front of the entrance slit to block any stray light entering into the

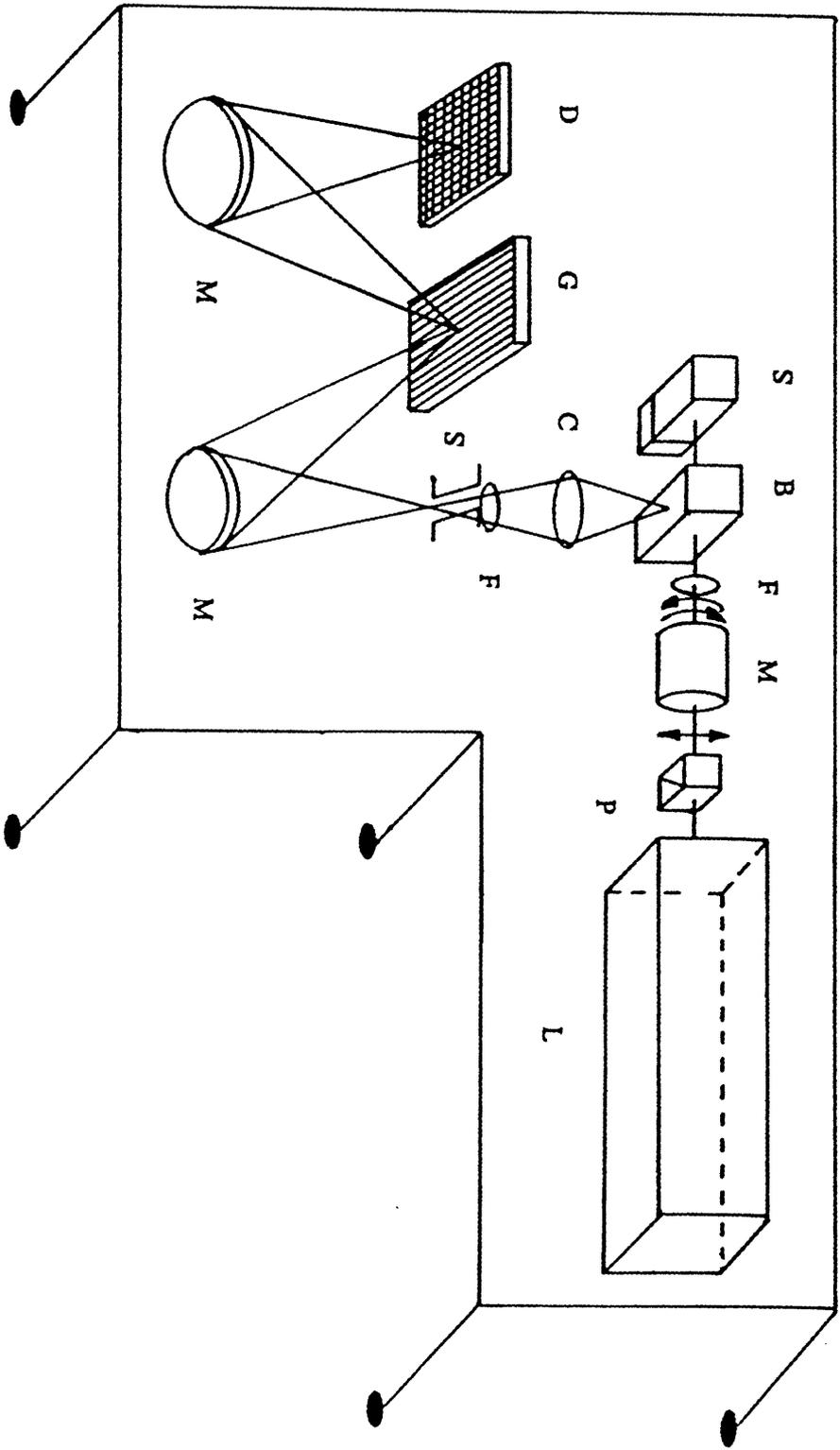


Fig. 3.1 A bird's eye view of the ROA Instrument.

spectrograph. The Raman signal is then dispersed across a grating (G) and exits onto the detector (D) attached at the exit port of the spectrograph. The Raman light is converted into electronic signal by the CCD detector and processed by computer.

### **3.3.1 Light source**

The laser used in ROA measurement is exactly the same as in a conventional laser Raman spectrometer but should take into account some additional considerations. Faint plasma lines accompanied with the major laser line were usually negligible in conventional Raman instrument but it could plague the ROA signal. Exciting wavelength and laser power available are two other factors which should be taken into account. The short wavelength exciting line generally gives rise to stronger ROA signal and high power is generally needed for better ROA sensitivity. We therefore chose the argon ion Stabilite 2016 (Spectra-Physics) laser as the light source in which two major emission lines 488.0nm and 514.5nm can be used. The 488nm line gives stronger Raman intensity and wider spectral coverage than the 514.5nm, but the latter generally produces less fluorescence background. It can be altered according to sample circumstance, which line would be preferred. We did try to employ a krypton ion laser for the expectation that it might considerably reduce the fluorescence background for biological sample. Unfortunately, two disadvantages are inherently associated with the krypton laser for ROA instrument. One is the loss of Raman intensity due to the Rayleigh scattering law. The other is that the krypton laser has large plasma lines

over almost the entire Raman spectrum region. Moreover, the fluorescence reduction by using krypton red line 647.1nm is not significant compared to the argon green line 514.5nm in many biological molecules.

### 3.3.2 Polarizer

A calcite Glan-Taylor polarizer (Leysop GT12) is set at the exit beam of the laser to ensure the perfect linear polarization of the laser beam before it enters into the modulator. This is essential because perfectly linearly polarized light is a prerequisite to generate perfectly circularly polarized light by the modulator.

### 3.3.3 Modulator

The optical polarization modulator generates the required circularly polarized light from the linearly polarized laser beam. It is a non-linear optical crystal  $KD^*P$  which is forced to produce  $+\lambda/4$  or  $-\lambda/4$  retardation when a characteristic positive or negative high voltage ( $\sim 3KV$ ) is applied upon it.<sup>16</sup> The modulator (Leysop Model 850-16) is specifically designed for ROA measurement in which the  $KD^*P$  crystal has a larger aperture of 16mm and an extinction ratios of  $2700 \pm 200:1$ . A longitudinal electric field is generated by gold ring electrode mounted on the faces of the  $KD^*P$  crystal and is driven by a high voltage linear differential amplifier (Leysop Model 5000). The amplifier features an individually adjustable positive and negative limits and when driven by a bipolar square wave, generates an ultrastable bipolar high voltage square wave for switching the initially

linearly polarized laser light alternatively between right and left circular polarization. The  $\text{KD}^*\text{P}$  crystal is temperature stabilized to within  $0.01^\circ\text{C}$  by using a Model TC 15 (Electro-Optic Developments) temperature controller. It is necessary to control precisely the temperature of the  $\text{KD}^*\text{P}$  crystal because its optical birefringence is a function of temperature. Any temperature drift will inevitably cause the deviation of the perfect circular polarization of the exit light from the modulator.

The modulator is mounted on a stage ( Micro-Controle) driven by a micrometer that has six degree of freedom ( three translation, two tilt and one rotation ). It allows precise alignment of the modulator within a resolution of  $\pm 1$  arc minute and  $1\mu$  displacement, respectively, and is used to align the modulator along the laser light emission direction and to rotate the crystal axes at  $45^\circ$  to the plane of linear polarization of the laser beam.

### 3.3.4 Focusing lens

The circularly polarized beam exited from the modulator is focused by a plano-convex focusing lens ( Melles Griot ) of 150mm focal length through the hole in the backscattering block into the sample. The aim of using a focused beam as Raman exciting light is to increase the power density on the sample in order to generate higher Raman photons because Raman intensity is proportional to the incident photon flux.<sup>17</sup> A focusing lens with short focal length is usually used in  $90^\circ$  scattering geometry to yield better power density than that of a lens with long focal length.<sup>17</sup> However, in backscattering arrangement, it is impractical to employ

focusing lens with very short focal length because the backscattering block is located between the sample cell and the laser focusing lens and there is no more space between them.

### 3.3.5 The backscattering block

The merit of backscattering geometry for ROA instrument has been discussed in the previous chapter 2. This is the central element in the optical train. The details of the block had been originally described by Hug<sup>18</sup> and a modified version is depicted in Fig. 3.2. A key element

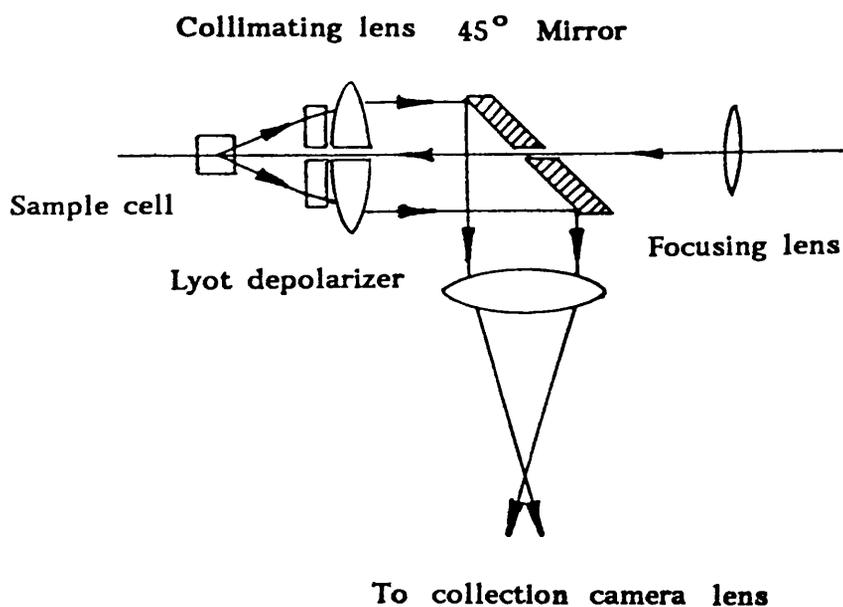


Fig. 3.2 The backscattering block.

composing the backscattering block is a Lyot depolarizer made up of two pieces of calcite crystal of 25mm diameter and 3 and 6mm thickness respectively (Leysop). It is mounted ahead of the achromatic triplet lens (Melles Griot) used to collimate the scattered Raman light from the scattering cone. The scattered polarized Raman photons are effectively depolarized before hitting the reflection mirror. The purpose of the depolarizer is to create an annular light collection aperture with equal efficiency for all solid angle  $\alpha$  regardless of the polarizing properties of the spectrograph.<sup>18</sup> Depolarized Raman light is then collimated onto the reflection mirror which direct the Raman light into the collection optics. An exit focusing lens of 38mm focal length (Optical Works Limited) is used to focus the reflected Raman light to a focal point to match the collection optics.

### 3.3.6 Sample cell

The sample cell is as important as any of the optical components constituting the whole optical train. It should be free of residual birefringence which is believed to be one of the sources of artifact. We used a quartz fluorescence cell (Optiglass) of  $5 \times 5 \times 10 \text{ mm}^3$  as the sample container. The actual solution volume used is about 100 $\mu$ l. A larger volume sample cell is favourable with respect to giving more Raman signal. However, it is unrealistic for biological molecules because they are usually only available in small quantity. The sample cell is held by a table type  $YZ\theta_y\theta_z$  mount (Micro-Controle) which provides four degree (two translation, one tilt and one rotation) of freedom of precision positioning

and thus ensures the perfect alignment of the sample cell in the optical train.

### **3.3.7 Collection optic and magnification**

The amount of scattered Raman photons that can be collected depends on the numerical aperture of the lens mounted on the backscattering block. The larger the numerical aperture the more scattered Raman light would be picked up. However, there are other factors that have to be taken into account as far as the total collection efficiency of the spectrometer is concerned. These include the magnification of the optical collection system and the match of the F/number between the spectrograph and the collection optic.<sup>19</sup> Large magnification has been exclusively employed in the conventional commercial Raman spectrometer which has usually a small acceptable solid angle. The total collection efficiency is thus dictated by the spectrographic aperture and spectral field. Small magnification or unity magnification of the collection optic is more favourable.

In the ROA instrument, after the Raman light exits from the backscattering block, a camera lens (Carl Zeiss Planar T) with 85mm focal length and F/1.4 aperture is used as the collection optic which matches well with the spectrograph (focal length 250mm, F/4.1). The actual magnification is arranged at about three and the spectrograph is slightly underfilling which can be learned from the image on the first mirror of the monochromator. The camera is mounted on a holder which can be smoothly adjusted in three dimensions via micrometers to obtain the

maximum Raman intensity.

### 3.3.8 Spectrograph

The heart of the ROA instrument is the monochromator which disperses the Raman light.<sup>19</sup> Previously, double or triple monochromators had to be employed to provide sufficient rejection of stray light in Raman instrument. However, a single grating monochromator will give much higher throughput, but the stray light is unacceptably large.

The advent of the holographic Bragg diffraction edge filter offers a new opportunity for single grating spectrographs in Raman spectroscopy since together they give rise to Raman spectrometer with a high throughput and excellent rejection of the Rayleigh stray light.

A single holographic grating monochromator HR-250S (Jobin-Yvon) is used as the Raman spectrograph. It contains a 1200 groove/mm holographic grating blazed at 630nm with a size of 58x58mm<sup>2</sup>. The mirrors and grating are arranged in a configuration of the asymmetric Czerny-Turner type. It features a high throughput of 58%, high spectral purity, is free of astigmatism and has a moderate resolution when it operates at fairly wide slit. The focal length of the monochromator is 250mm and the spectral aperture is F/4.1. Its entrance slit can be adjusted manually within the width 0.5~2mm and height from 2 to 8mm. The monochromator is also compact, robust and very simple to be operated and convenient to be aligned to match the rest of the optical train.

### **3.3.9 Holographic Bragg diffraction filter**

The holographic edge filter is made up of precisely laminated polymer films by holographic technique.<sup>20</sup> The principle of a holographic Bragg diffraction filter is that only those light wavelengths fulfilling the Bragg diffraction condition will be diffracted while light with other wavelengths will be transmitted. In this way, the filter diffracts at the wavelength predetermined in the fabrication. Rayleigh stray light at given wavelength is diffracted and the Raman photons are allowed to transmit through the filter since their wavelength do not fulfill the Bragg diffraction condition.

In the ROA instrument a holographic Bragg diffraction (HBD) filter (25mm diameter, Physical-Optics Corporation) is set as close as possible to the front entrance slit of the monochromator. It has an optical density better than 5 and a 80% transmission over almost the whole Raman spectrum range and can be tuned in an incidence angle range of 5-20 degree. It shows that stray light within  $300\text{cm}^{-1}$  from the Rayleigh scattering line has been completely removed and leaves an extremely low background for the Raman spectrum. ( This edge filter has now been replaced with an even better notch filter which has 90% transmission, an O.D. over 6, a better rejection of stray light and a lower wavelength cut-off  $\sim 200\text{cm}^{-1}$ ).

### **3.3.10 CCD detector**

The last component in the optical train of the ROA instrument is the cooled CCD detector which is a new generation light detector based on

silicon photon-electronics technology. The main advantages of a CCD detector over previous multichannel light detectors such as the photo-diode array in Raman spectroscopy are low readout noise and high quantum efficiency which yields the high sensitivity desired by ROA measurement.<sup>21, 22</sup>

The CCD is made up of silicon semiconductor. The intrinsic quantum efficiency of all silicon detectors is high compared with that of available photocathode materials. Quantum efficiency losses due to absorption are greatly reduced by a backthinned design which thins the integrated circuit wafer thickness to the range of  $2\mu$  to  $20\mu$  by a chemical etching process with illumination from the back instead of the front. Backthinned CCDs also reduce the detection of cosmic ray, high energy particles and radioactivity events and therefore improves the SNR. Anti-reflection coating is used in CCD to reduce reflection losses and backthinned CCD is particularly amenable to anti-reflection coating. The charge collection efficiency of CCD is approaching 100%. The overall quantum efficiency of a typical backthinned CCD thus exceeds 90% in the blue. Moreover, the dark current, which is generated thermally, for cooled CCD is in the range of  $<0.001$  to  $0.003$  electrons per detector element per second which is almost negligible. The employment of CCD in Raman spectroscopy can further benefit by the on-chip binning technique to improve SNR. Other superiorities include linearity, geometric stability, large dynamic response range over ten orders of magnitude which is a remarkable feature uncompetitive by other detectors and robust to over-exposure, which is particular a weakness in the previous generation of photo-diode array

### Quantum Efficiency (%)

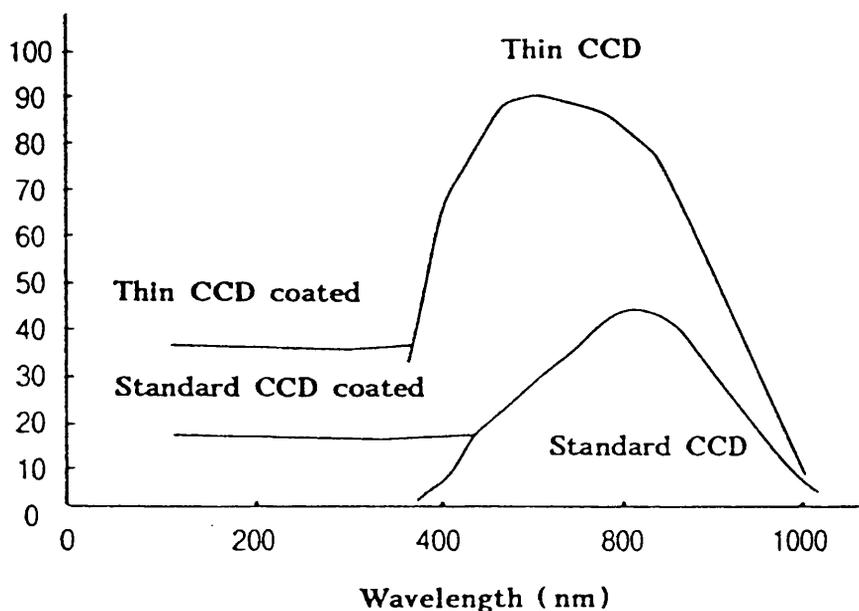


Fig. 3.3 Quantum efficiency curves of typical CCD detector in the UV-visible region.

detectors. It can be said that backthinned CCD detector is almost an ideal detector for ROA.

A Wright Instruments Ltd backthinned CCD detector is used in our ROA instrument. This is composed of a EEV P86231/T CCD chip with 385x578 pixels, and a camera electronic unit and a computer interface board. It is cooled by a four stage Peltier cooler to maintain an operating temperature at 200°K to reduce dark current below 0.1 electrons/pixel/second. Its peak quantum efficiency is over 80% at 500nm. Fig. 3.3 shows quantum efficiency of a standard and a backthinned CCD in

the whole visible region. The detector pixels can be binned on chip in vertical direction to reduce readout noise. The maximum charge storage capacity is around 300,000 ADC counts/pixel in low speed readout mode. The Raman light is dispersed across the long edge of the CCD chip to cover wider spectral range. The spectral coverage is governed by the exciting laser line, spectrograph grating grooves/mm and the total pixels on the long edge of the CCD chip. With argon ion laser line 514.5nm and 1200g/mm grating, the long edge of the CCD chip can cover approximately  $1150\text{cm}^{-1}$  spectral region. All the pixels are interpolated and it means that a pixel covers three wavenumber.

### **3.4. Electronic and computer system**

The schematic flow process of the electronic signal of the ROA instrument is depicted in Fig. 3.4. It involves two basic functions that control the synchronization of the switching of the polarization modulator and the electronic shutter of the CCD camera and register the Raman signal from the CCD detector.

#### **3.4.1 CCD electronics**

The operation of the CCD detector is controlled by an electronic module interfaced with the computer which contains three circuit boards. A baseboard contains power supplies and circuitry to handle the serial data link to the computer interface. An auxiliary board contains the power supply and temperature regulating control circuitry for the Peltier cooler

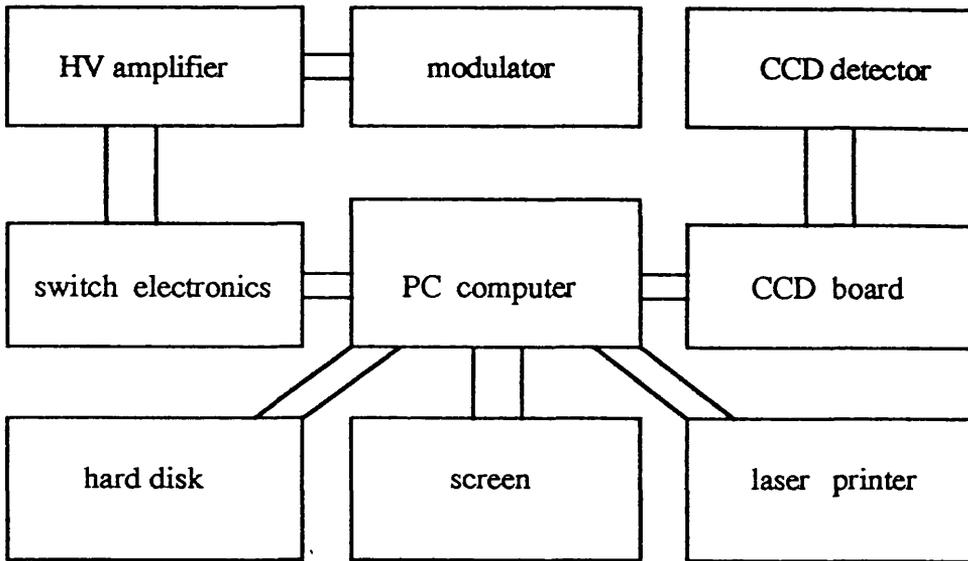


Fig. 3.4 The flow chart of the electronic signal of the ROA instrument.

and circuitry to control the mechanical shutter of the camera. The CCD board comprises all the circuitry to bias and clock the CCD and to amplify, sample and digitize the output signals. A MK.2 board offers multiple operations for each command received, together with faster clocking operations. It provides two different readout rates and a 16 bit ADC. The data link connecting the camera to the computer interface is bi-directional and operates at a data rate of 10 Mbaud.

### **3.4.2 Polarization modulation control**

The circular polarization modulator is synchronized with the registration of Raman spectra via an auto programmed module. In this way, the ROA experiment is performed automatically by the computer. The module is able to be pre-programmed with a complicated acquisition sequence, including the generation of a  $1\mu\text{s}$  TTL (Transistor-Transistor Logic) pulse at the start of each cycle. These pulses are fed into an electronic circuit which generates two synchronized output signals: with successive TTL pulses, one output supplies alternate positive and negative voltage levels to the high voltage linear differential amplifier to switch the polarization of the laser beam between right and left circular; the other output supplies alternate high and low TTL levels to the detector board to control the switch of the mechanical shutter and the sense of the acquisition (i.e. to add to obtain Raman spectrum or subtract to yield ROA). The electronic module also ensures that the modulation cycling always starts with the right phase of the modulation and stops the modulation cycle at a certain time (40s) once a ROA experiment has been accomplished and returns the voltage applied to the modulator crystal to zero in order to protect it from damage by the high D.C. voltage. The output signals toggle with each input TTL pulse from the module, thus allowing to accumulate the ROA difference spectrum directly.

### **3.4.3 Computer and software**

The ROA instrument is fully controlled by a high-speed, high

performance AT personal computer ( Dell 325D ). The computer features an Intel 80386 microprocessor running at a clock speed of 25 MHz and plus an Intel 80387 math coprocessor, a 1 megabyte system board memory, a 32 kilobyte RAM (ramdom access memory), a 99 megabyte hard disk drive and a 1.2 megabyte floppy diskette drive. A MS-DOS and a Micro GW-basic interpreter has been used as the computer operation system.

The acquisition of ROA spectra and the synchronization of the CCD detector and polarization modulator is performed by an interface board manufactured by Wright Instruments installed on the computer. A VGA Color Plus monitor (Dell) and a laser printer (Hewlett-Packard) are employed as spectrum display and presentation.

A Lab Calc software package specifically designed for laboratory data acquisition and manipulation (Galactic Industries Corporation ) is employed as the ROA instrument operation software and for post-experiment data manipulation. A ROA subroutine is incorporated into the original Lab Calc system by Wright Instruments for ROA data acquisition, instrument configuration and spectrum data manipulation.

#### **3.4.4 ROA spectrum acquisition**

ROA measurement is executed by a single touch driving-menu, once the sample cell alignment is achieved. The high voltage amplifier is set to external to receive command from computer to switch the polarization of the modulator. The electronic box for switching of the polarization is reset to zero to ensure that the amplifier is ready to receive the TTL

pulse. The computer sends order to the high voltage amplifier to output the high voltage square wave which force the modulator to change birefringence and transform the linearly polarized incident laser light into circularly polarized light (right); simultaneously, the mechanic-electronic shutter of the CCD detector opens and the CCD chip receives Raman light and stores the charges converted from the Raman signal.

A preset exposure time of the CCD detector is determined by the sample circumstance in order to obtain the maximum SNR and to minimize the dead time. Once the preset single exposure time (it is usually set with less than 10 seconds because longer exposure time may damage the modulator) has expired, the shutter closes and a readout process is performed on the CCD detector and the charges are transferred to the on-chip amplifier and digitized and sent to the computer. A Lab-Calc memory 1 is used to register the first cycle of the data which is written as Raman intensity of right circular polarization. The CCD chip is then cleared and the software sends a second TTL pulse to the high voltage amplifier to change the phase of the high voltage square wave and the modulation polarization (left) and command the CCD shutter to open again. The process is continued as before and the Raman signal is recorded in a second memory 2. The Raman intensity difference is obtained by subtracting signal from memory 1 by 2 and stored in memory 3. The memory 1 and 2 are cleared and the acquisition is cycled and the Raman intensity difference (ROA) spectrum are accumulated in memory 3 until a preset time is reached. The Raman summation intensity is simply the sum data in memory 1 and 2 multiplied by the pair number of acquisition. The

raw spectra of Raman and ROA are then stored in the hard disk and can be further manipulated by the Lab-Calc system. However, raw data form are usually presented without any smoothing.

### 3.5. Spectrum calibration

The Raman frequency calibration is achieved by using a Raman standard calibration procedure with a known Raman frequency sample. For the sake of convenience, we choose  $\alpha$ -pinene, which is also extensively used as an ROA model sample, and indene, which has abundant well resolved Raman bands in the wanted frequency region, for the relative Raman frequency calibration.

The dispersion of the spectrometer is calculated by reading the known sample Raman spectrum on screen in the raw chip frame which is marked with pixel number at fixed grating scanning wavelength  $\lambda$ . The dispersion is obtained as follow:

$$D_{\text{nm/pixel}} = \frac{(\nu_a - \nu_b)}{(P_a - P_b)} \times \frac{1}{10000} \quad 3.1.$$

where  $\nu_a$ ,  $\nu_b$  are two known Raman bands frequencies centered at the pixel number  $P_a$ ,  $P_b$  respectively. It is usually chosen that the two bands are located near the two edges of the CCD chip respectively. Once the dispersion is determined, the Raman frequency is calibrated according to the following equation

$$\Delta\nu = 10^7 \left\{ \frac{1}{\lambda_o} - \left( \frac{1}{(c-a)d + \lambda_c} \right) \right\} \quad 3.2$$

where  $\lambda_o$  is the exciting laser line wavelength,  $\lambda_c$  is the wavelength scanned at the spectrograph,  $c$  is the pixel reading corresponding to the  $\lambda_c$  reading,  $a$  is the actual pixel and  $d$  is the dispersion in nm/pixel across the CCD chip. This calibration can yield a precision of  $\pm 3\text{cm}^{-1}$  and is equivalent to the procedure recommended by Hamaguchi<sup>23</sup> for multichannel spectrometer calibration.

## **3.6 Experiment details**

### **3.6.1 Sample handling**

Sample preparation constitutes an important link in the success of ROA measurement. It is more demanding for sample handling in ROA than in conventional Raman. Organic neat liquids are generally distilled by a specially designed mini-distillation device to purify the sample and reduce fluorescence background. For biological molecules in aqueous solution, high purity of the sample is a prerequisite. Samples are generally dissolved in distilled water and filtered through  $0.22\mu$  Millipore micro pore membrane filter to remove any particles with larger or equivalent size to the exciting light wavelength to avoid Tyndall scattering. Further centrifugation is generally required to ensure no particles and voids are present. Charcoal treatment was found to be very effective to reduce fluorescence background for small biological molecules such as monosaccharides. For biological macromolecules, purification of the sample via conventional physico-chemical methods are essential. The cleanness of the inner and outer sample cell wall is ensured by chromic acid and

washed thoroughly with distilled water and rinsed with filtered methanol.

### 3.6.2 Artifact control

The elimination of artifacts is a prerequisite to record reliable ROA spectra, especially when strongly polarized Raman bands are present in the Raman spectrum of the sample studied. As mentioned before, the strategy to reduce artifacts involves multiple efforts. First, the electro-optic modulator is carefully aligned and the quarter-wave voltages set accurately by using a Babinet-Soleil compensator. Second, the modulator is temperature stabilized by a high stable temperature controller with an ultra stable output voltage supplied by the linear differential amplifier that ensures a high degree of stability of the modulator. Third, the use of uncoated fused quartz throughout the optical train to minimize strain birefringence. The last but not least component to control artifacts is a Lyot depolarizer, which depolarizes the Raman light cone before hitting on the reflection mirror. In practice, it is rotated manually to locate a proper orientation at which the Lyot depolarizer has the maximum depolarization efficiency and minimize the artifact of the strongest polarized band as a criterion. The strong polarized Raman bands at  $667\text{cm}^{-1}$  in  $\alpha$ -pinene and at  $875\text{cm}^{-1}$  in the di-L-alanine are usually used to adjust the Lyot orientation for organic liquid and aqueous solution respectively. The Lyot position is then fixed once the strongest polarized band shows minimum ROA artifact. Very small offset voltage on the differential amplifier is usually applied to fine control the artifact.

### 3.6.3 Stray light and base line

Stray light is another obstacle to obtaining good quality ROA, especially at lower frequency. It becomes more severe in biological molecules in aqueous solution and increases with the molecular weight of the biological macromolecules and concentration. Practically, the holographic Bragg diffraction edge filter is set as close as possible to the entrance slit and tilted to a favourable angle to suppress the Rayleigh line according to sample circumstance by monitoring the Raman bands intensities at lower frequency.

Baseline in a ROA spectrum refers to the balance between positive and negative parts. It is influenced by both the stray light at lower frequency side and fluorescence at the higher. Poor alignment of the sample cell position and wrong Lyot depolarizer orientation may also distort the balance. It seems that a proper adjustment of the collection optic is essential to a balanced baseline. This is usually achieved by monitoring the two dimensional image of the Raman spectrum, while precisely adjusting the horizontal coordinate of the collection camera via micrometer. A nearly perfect symmetric two dimension image of the Raman spectrum is practiced as a criterion. Again, very tiny voltage offsets on the high voltage amplifier are usually employed to compensate a small drift of the baseline during long time aquisition.

### 3.7. The performance of the ROA instrument

#### 3.7.1 Instrument sensitivity

The remarkable sensitivity of the ROA spectrograph is exemplified by the ROA spectrum of  $\alpha$ -pinane taken in just 10 second (Fig. 3.5). All its main ROA features appear. Fig. 3.6 shows the ROA spectrum run for 10 minutes which indicates that its sensitivity is sufficient for rapid routine measurement on pure organic liquids.

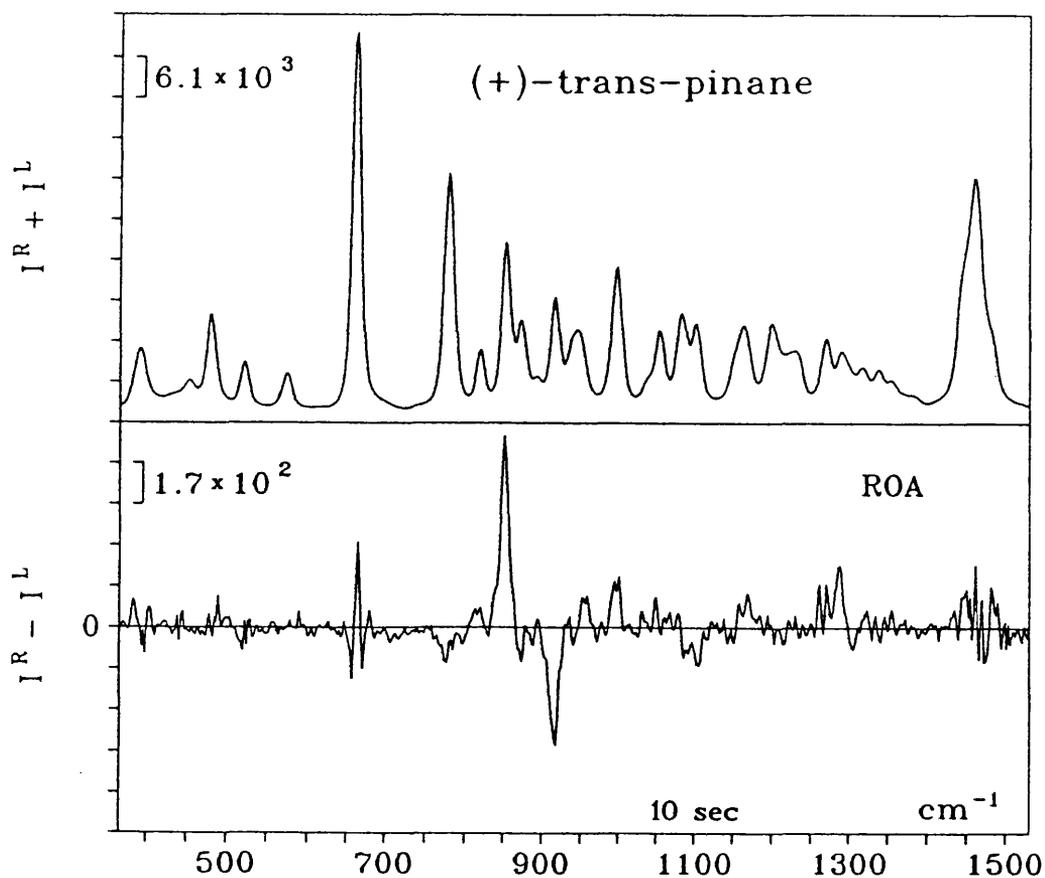


Fig. 3.5 Raman and ROA spectra of (+)-trans-pinane taken at 10 seconds.

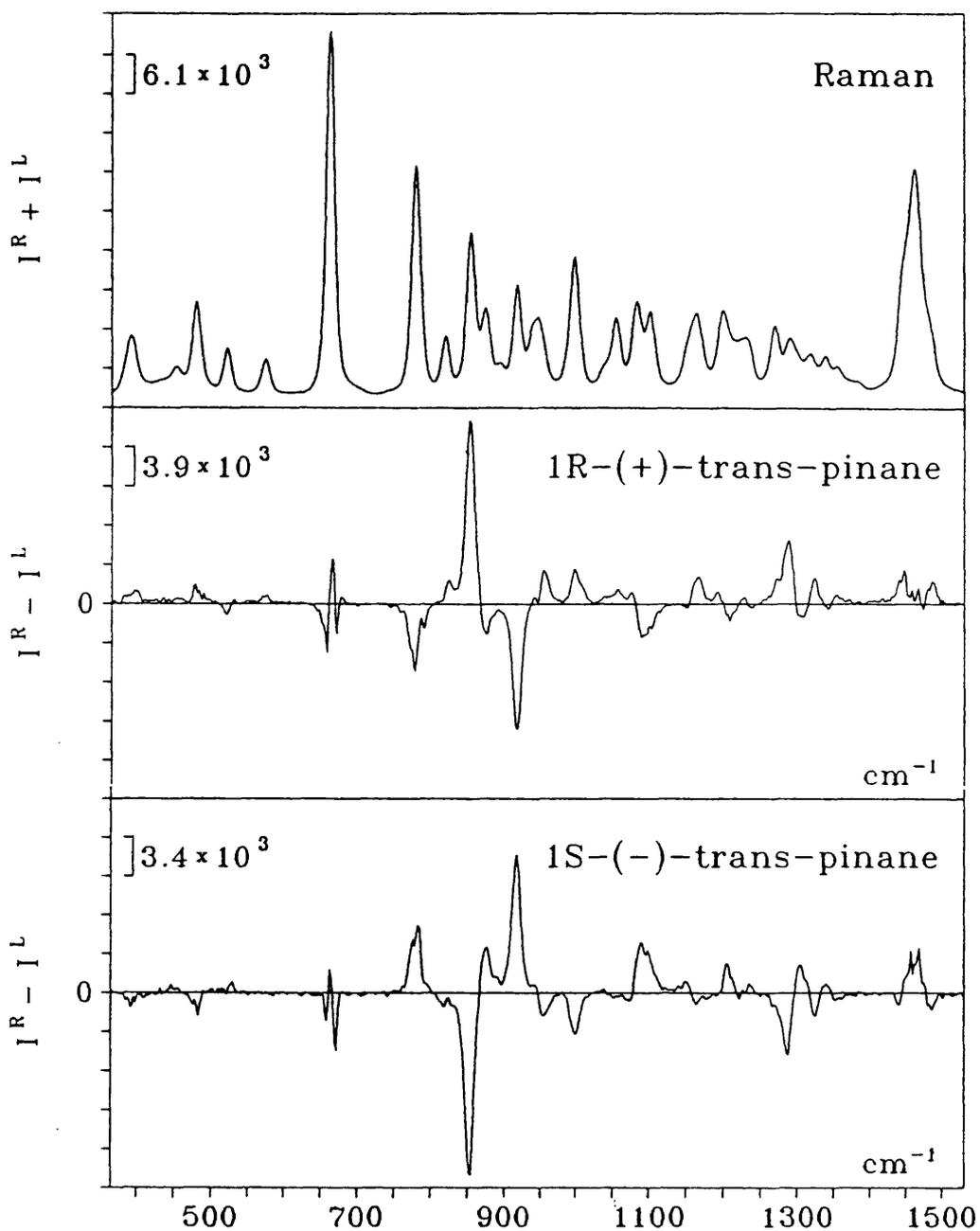


Fig. 3.6 Raman and ROA spectra of a pair of enantiomers 1R-(+)-, 1S-(-)-trans-pinane in neat liquids.

For biological molecules in aqueous solution, even higher instrument sensitivity would be desirable. Nevertheless, good quality ROA spectra can be recorded within one hour for carbohydrates, a few hours for small peptides and nucleic acids and a working day for routine proteins. Fig. 3.7-3.9 present ROA spectra of some simple biological molecules D-glucose, di-L-alanine and its enantiomer di-D-alanine, and  $\alpha$ -D-cyclodextrin.

The improvement of ROA instrument sensitivity for protein measurement is best demonstrated by the comparison of the first segment of ROA spectrum of a protein lysozyme, which was recorded on an old

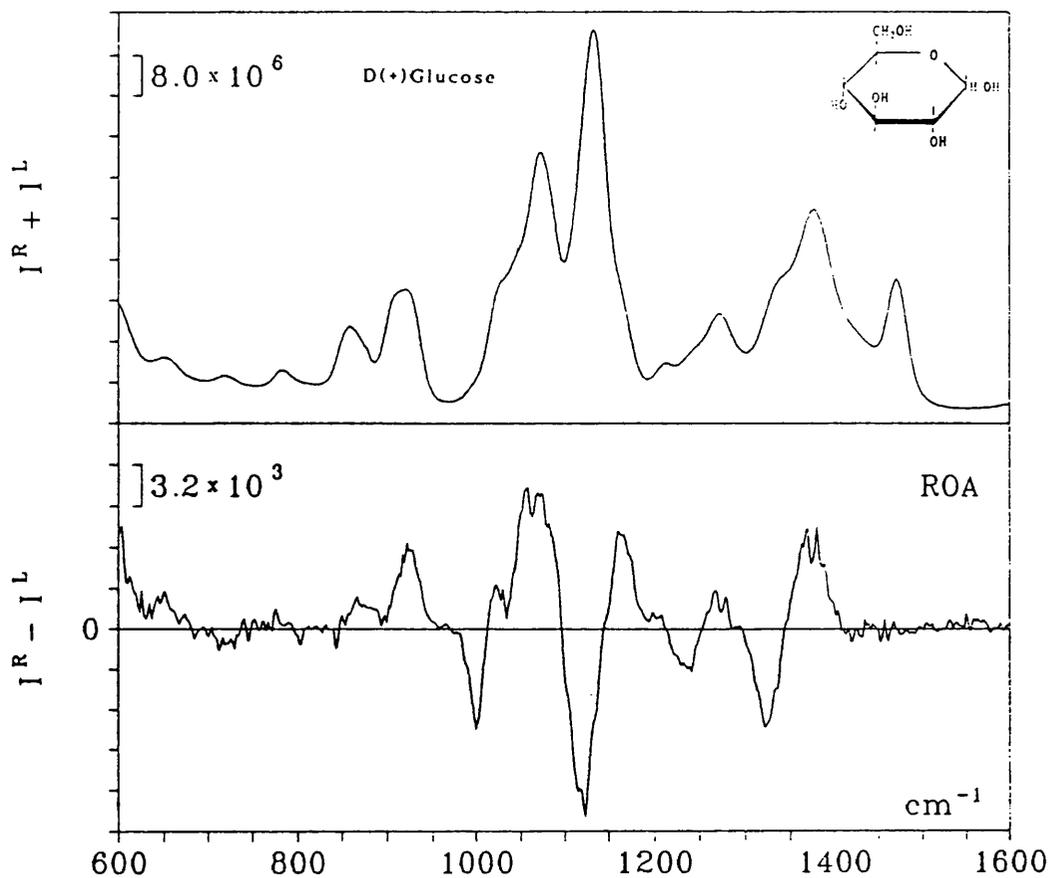


Fig. 3.7 Raman and ROA spectra of D-glucose in water.

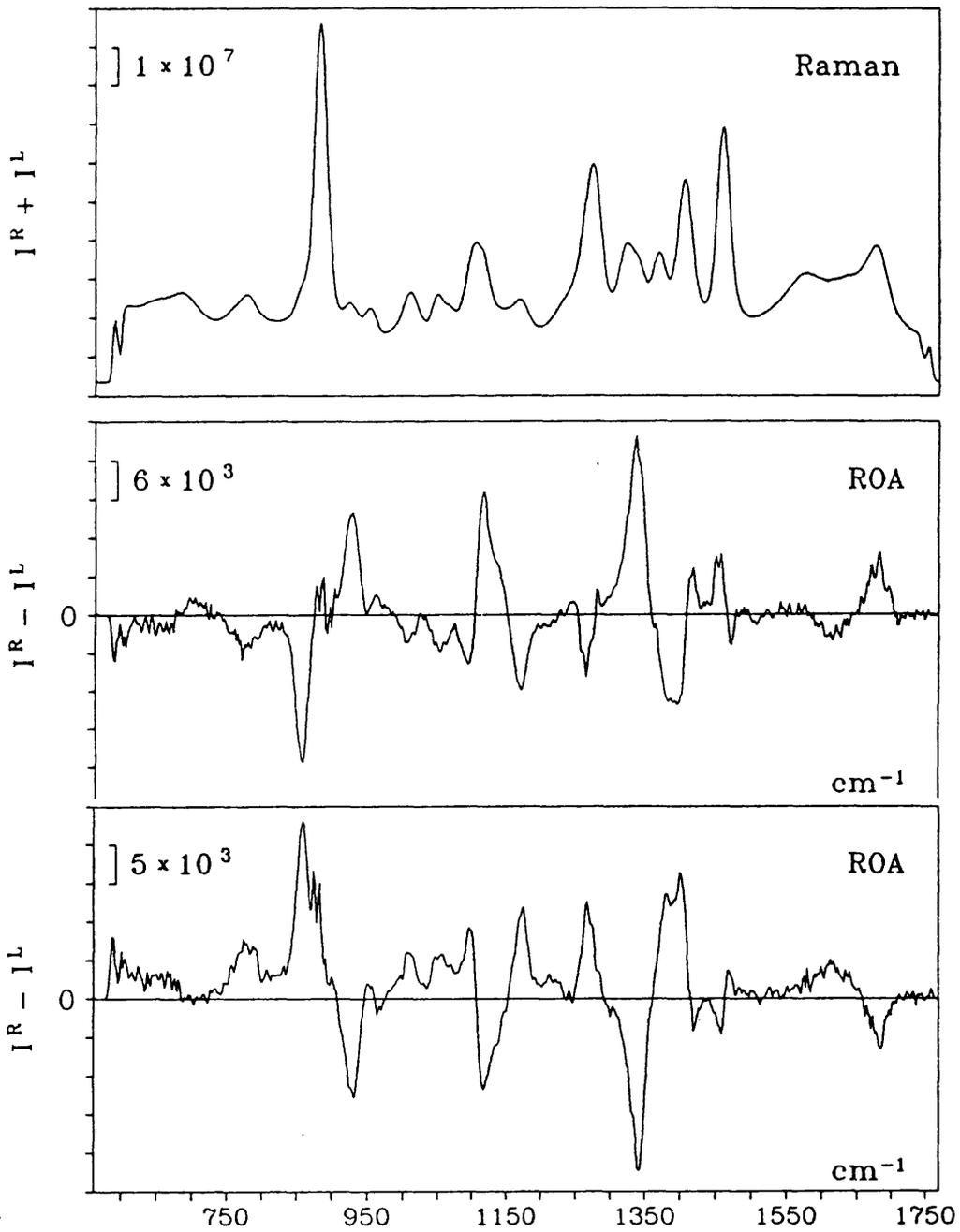


Fig. 3.8 Raman and ROA spectra of di-L-, D-alanine in water.

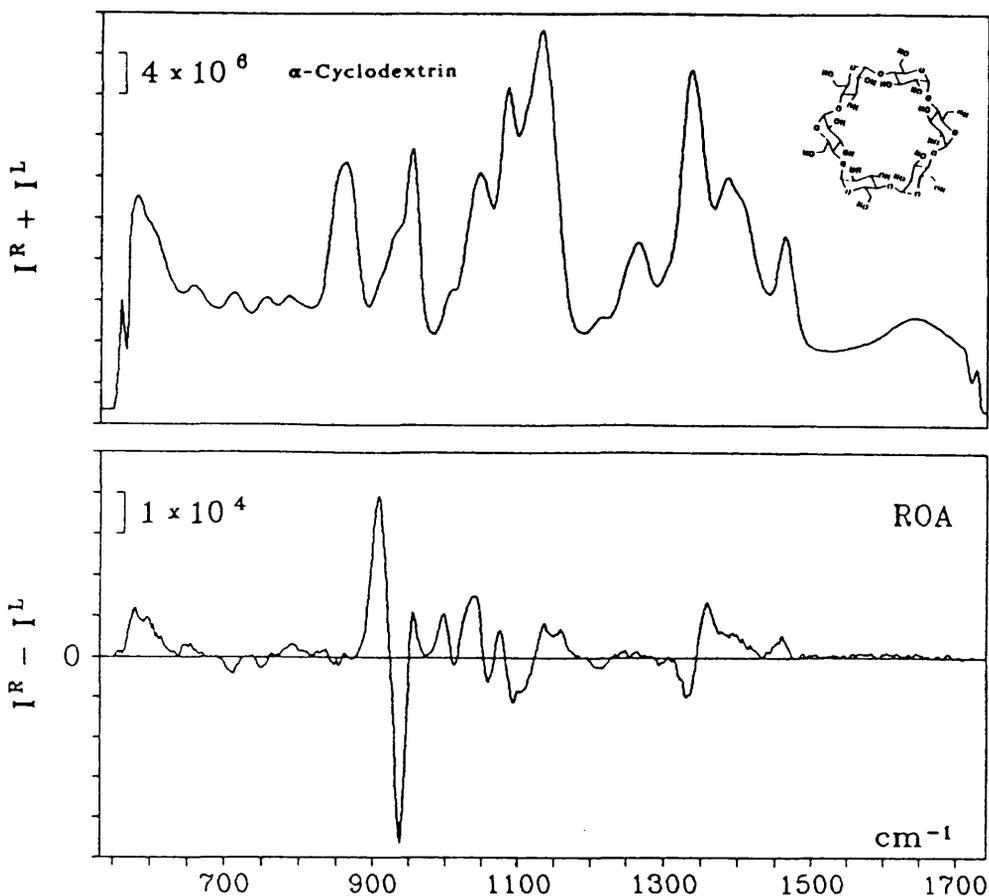


Fig. 3.9 Raman and ROA spectra of  $\alpha$ -D-cyclodextrin in water.

double monochromator spectrograph with an ordinary CCD detector, with the one obtained on the new ROA instrument (Fig. 3.10). A rough estimate of the SNR and the measurement time reveals that approximately 100 fold more measurement speed has been achieved with the new generation ROA instrument compared to the second generation ROA instrument based on a multichannel photo diode array detector and a  $10^5$  times instrument sensitivity has been achieved since the first measurement of ROA spectrum on a single scanning Raman spectrometer two decades ago. Further improvement of ROA instrument sensitivity is anticipated with the

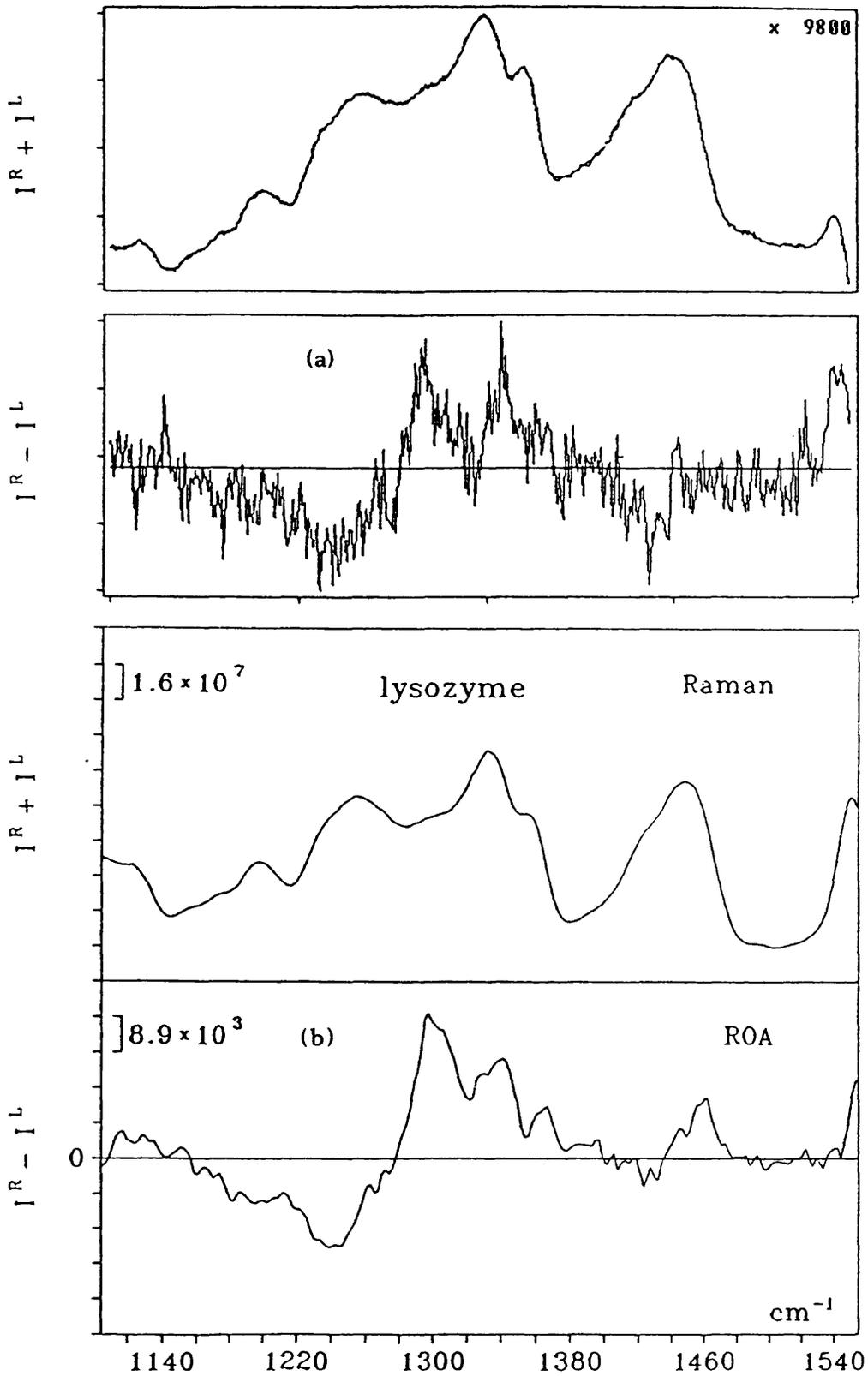


Fig. 3.10 A comparison of instrument sensitivity for proteins. (a), lysozyme recorded on the old instrument, (b), recorded on the new instrument.

improvement of the collection optic by using a non-imaging collector<sup>24</sup> and the new development of spectroscopic technology, for instance, a faster spectrograph.

### 3.7.2 Spectral resolution

The ROA spectrum resolution is closely related to the instrument sensitivity. At present, coarse resolution is practiced in order to trade for better instrument sensitivity. Generally, a slit width of  $120\mu$  is practiced for biological molecules and a narrower slit can be selected for pure organic liquids. The ROA spectrum recorded at an entrance slit of  $120\mu$  shows approximately a  $12\text{cm}^{-1}$  resolution, which is fairly acceptable because the band width of most Raman bands for biological molecules in aqueous solution is  $\sim 10\text{-}20\text{cm}^{-1}$ .

### 3.7.3 Spectral coverage

The spectral coverage of the ROA instrument depends on the wavelength of the exciting light, the pixels available at the long edge of the CCD detector and the grating groove density. With  $514.5\text{nm}$  argon line,  $1200\text{g/mm}$  grating and 578 pixels in the dispersion direction of the CCD detector, the spectral coverage of the ROA instrument for a single exposure is  $\sim 1150\text{cm}^{-1}$ . Obviously, any spectral range from  $200\sim 3500\text{cm}^{-1}$  can be measured by simply scanning the grating to the central wavelength of interest and running a second spectrum. The two spectra can be combined by the slice function of the data manipulation software.

## REFERENCES

1. L. D. Barron and J. Vrbancich, *Top. Cur. Chem.*, **123**, 151, 1984.
2. H. Boucher, T. R. Brocki, M. Moskovits and B. Bosnich, *J. Am. Chem. Soc.*, **99**, 6870, 1977.
3. W. Hug and H. Surback, *Chem. Phys. Lett.*, **60**, 186, 1979.  
Wiley-Heyden, New York, 1, 1982.
4. L. D. Barron, J. F. Torrance and J. Cutler, *J. Raman Spectrosc.*, **18**, 281, 1987.
5. L. D. Barron, L. Hecht, W. Hug and M. J. Macintosh, *J. Am. Chem. Soc.*, **111**, 8731, 1989.
6. L. Hecht, L. D. Barron and W. Hug, *Chem. Phys. Lett.*, **158**, 341, 1989.
7. L. Hecht and L. D. Barron, *Appl. Spectrosc.*, **44**, 483, 1990.
8. L. Hecht, D. Che and L. A. Nafie, *Appl. Spectrosc.*, **45**, 18, 1991.
9. D. Che, L. Hecht and L. A. Nafie, *Chem. Phys. Lett.*, **180**, 182, 1991.
10. L. Hecht, L. D. Barron, A. R. Gargaro, Z. Q. Wen and W. Hug, *J. Raman Spectrosc.*, **23**, 401, 1992.
11. D. S. Kliger, J. W. Lewis and C. E. Randall, *Polarized Light in Optics and Spectroscopy*, Academic Press, New York, 1990.
12. W. Hug, *Appl. Spectrosc.*, **35**, 115, 1981.
13. L. D. Barron and J. Vrbancich, *J. Raman Spectrosc.*, **15**, 47, 1984.
14. L. Hecht, B. Jordanov and B. Schrader, *Appl. Spectrosc.*, **41**, 295, 1987.
15. P. L. Polavarapu, *Spectrochim. Acta*, **46A**, 171, 1990.
16. J. M. Ley and P. J. Webb, *Electronics Letters*, **4**, 213, 1968.
17. J. J. Barrett and N. I. Adams, *J. Opt. Soc. Am.*, **57**, 19, 1967.

18. W. Hug, in *Raman Spectroscopy*, edited by J. Lascombe and P. V. Huong, Wiley -Heyden, New York, 1, 1982.
19. J. M. Lerner and A. Thevenon, *The Optics of Spectroscopy*, Jobin-Yvon, 1990.
20. M. J. Pelletier and R. C. Reeder, *Appl. Spectrosc.*, **45**, 765, 1991.
21. R. B. Bilhorn, J. V. Sweedler, P. M. Epperson and M. B. Denton, *Appl. Spectrosc.*, **41**, 1114, 1987.
22. R. B. Bilhorn, P. M. Epperson, J. V. Sweedler and M. B. Denton, *Appl. Spectrosc.*, **41**, 1125, 1987.
23. H. O. Hamaguchi, *Appl. Spectrosc. Rev.*, **24**, 137, 1987.
24. W. T. Welford and R. Winston, *High Collection Nonimaging Optics*, Academic Press, London, 1989.

## CHAPTER 4

# Raman optical activity of peptides and polypeptides

### 4.1. Introduction

Peptides are the basic building blocks of proteins. They are composed of a small number of amino acid residues. Peptides have also been recognized as hormones, growth factors, antibiotics, antitoxins, ionophores and are involved in a wide range of biological functions. There is increasing interest in developing peptides and oligopeptides as new therapeutic agents in the pharmaceutical industry. It is therefore important to study the structure and conformation of peptides and to establish the correlation between their structure and bioactivity.

Globular proteins have relatively short secondary structural segments: for instance,  $\alpha$ -helical components are usually not very long (three or four turns); the largest  $\beta$ -strand includes only about 15 residues and reverse turns are generally associated with three or four residues. The biologically active sites in a protein are usually associated with certain peptide segment which form a specific local environment suitable for ligand binding or enzyme activity. Structural information on equivalent peptide segments is particularly important for understanding the protein folding and functions. The study of peptide structure and conformation can thus be used for dual purposes. On the one hand, they are model compounds for proteins, and on the other, to determine the structure and

conformation of peptides themselves is essential to elucidate their biological functions and interactions with other molecules.

Raman spectroscopy has played an important role in determining structure and conformation of peptides and polypeptides together with many other spectroscopic techniques such as ECD and NMR. The merit of Raman spectroscopy is that certain vibrational modes of the peptide linkage, in particular, amide I and amide III, exhibit characteristic band frequencies which have been correlated with the secondary structure or chain backbone conformation of peptides and polypeptides.<sup>1</sup> The study of Raman spectra can thus provide the information on the chain conformation. Moreover, Raman spectra can be recorded in aqueous solution, which is the native environment for many peptides and proteins. With its enhanced resolution power and sensitivity, Raman optical activity measurements should provide much more subtle information on peptides and polypeptides.

In this chapter, detailed ROA study is focused on a number of model peptides; in particular the alanine oligomers di-L-alanine, tri-L-alanine and tetra-L-alanine in various aqueous solution environment ( H<sub>2</sub>O, D<sub>2</sub>O, HCl, NaOH ) and a tripeptide, L-Pro-L-Leu-Gly-amide, which is a model for the  $\beta$ -turn structure that has been recognized to play a crucial role in polypeptide chain folding in globular proteins. Two polyamino acids, poly-L-lysine and poly-L-glutamic acid are also included since they are basic polypeptide chain models for proteins. This chapter also forms the springboard for further ROA studies on proteins which is the main theme of the following chapter 5.

## 4.2 Raman optical activity of di-L-alanine

Di-L-alanine is a dipeptide consisting of two L-alanine residues joined by one peptide linkage. Its molecular frame is depicted in Fig. 4.1. This is an ideal molecule for the initial ROA study of peptides in aqueous solution for a number of reasons. Di-L-alanine is one of the simplest peptides and its structure and conformation have been extensively investigated by vibrational spectroscopy including both Raman and infrared.<sup>2</sup> It has been used as a prototype molecule in conventional Raman studies of alanine oligomers.<sup>2</sup> X-ray diffraction study has revealed that it adopts an extended chain conformation in crystals,<sup>3</sup> The solution

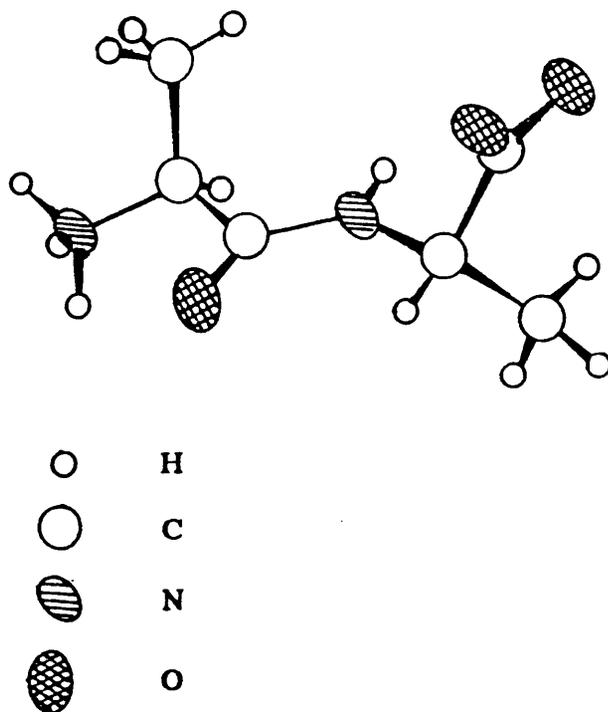


Fig. 4.1 The schematic view of the molecular conformation of di-L-alanine.

conformation of di-L-alanine is suggested by NMR measurement to be very similar to that in the crystal.<sup>4</sup> Furthermore, ROA measurements in conjunction with theoretical *ab initio* ROA calculations on the amino acid L-alanine provided a basis for the analysis of ROA spectra of alanyl peptides.<sup>5</sup>

It is well known that conventional vibrational spectra analysis of peptides and proteins via Raman and infrared relied heavily on the utilization of the model compound N-methylacetamide in the early days. It was the detailed vibrational analysis of this model compound that established the primary relationship between the vibrational band frequency and the peptide linkage of the polypeptide chain and laid the foundation of the polypeptide and protein conformation study by vibrational spectra.<sup>6</sup> However, N-methylacetamide is not a suitable model compound for vibrational optical activity study of peptide simply because it is not chiral. Moreover, recent detailed vibrational spectra study of alanyl-alanine dipeptides have found that the previous assignment of the amide III mode of the peptide linkage based on the vibrational analysis of the N-methylacetamide is not sufficient.<sup>7</sup> Vibrational optical activity via VCD has been employed to study the alanyl-alanine dipeptides in aqueous solution to help the explanation of the conformational sensitivity of the amide III modes.<sup>8</sup> It is therefore considered that the di-L-alanine dipeptide could be used as a better model peptide for both vibrational spectra and particularly vibrational optical activity analysis of polypeptides and proteins.

ROA measurements on di-L-alanine were carried out in different

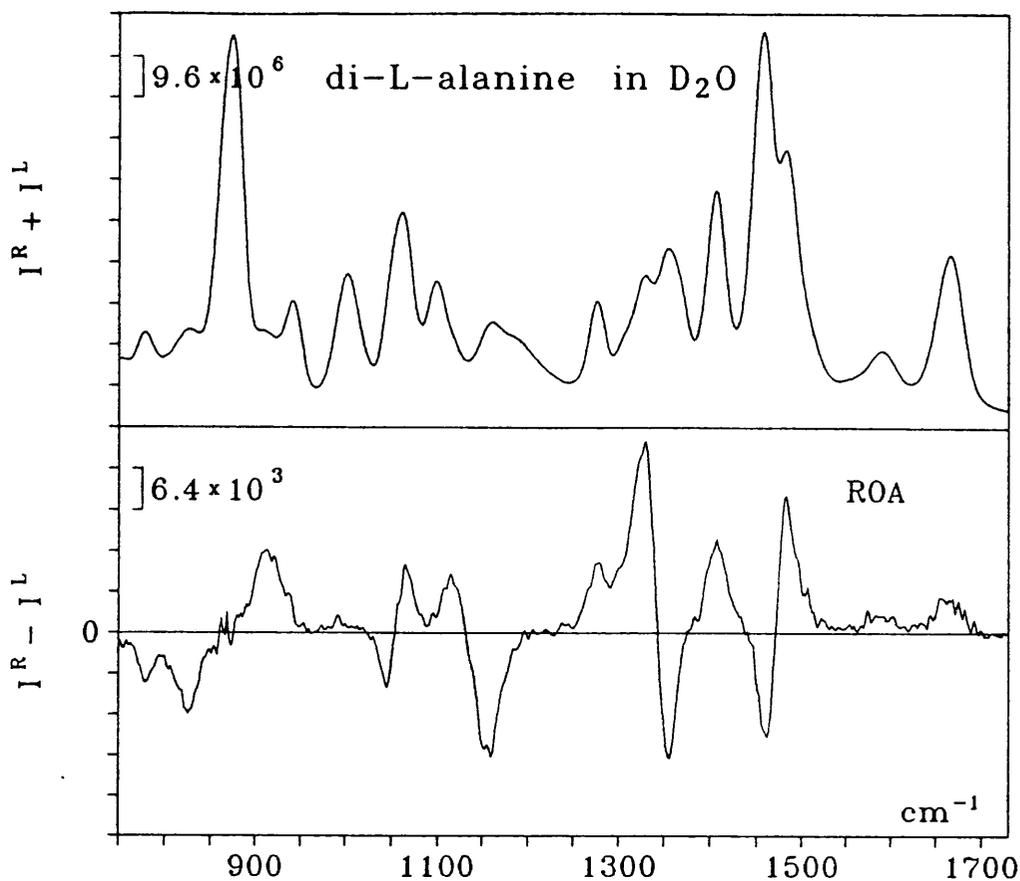


Fig. 4.2 Raman and ROA spectra of di-L-alanine in  $D_2O$ .

aqueous solution, e.g. in water and  $D_2O$  with neutral pH and base and acid pH (5N NaOH and 5N HCl). The ROA spectrum of the enantiomer di-D-alanine in water was also measured to ensure the reliability of the data obtained from the new instrument.

The Raman and ROA spectra of the di-D-alanine and di-L-alanine have been presented in chapter 3 (see Fig. 3.8 ). Fig. 4.2 shows the Raman and ROA spectra of di-L-alanine in the  $1730-750\text{cm}^{-1}$  region in  $D_2O$ . Fig. 4.3 and Fig. 4.4 show the Raman and ROA spectra of di-L-alanine in 5N HCl and 5N NaOH respectively. The ROA spectrum below  $750\text{cm}^{-1}$  were not

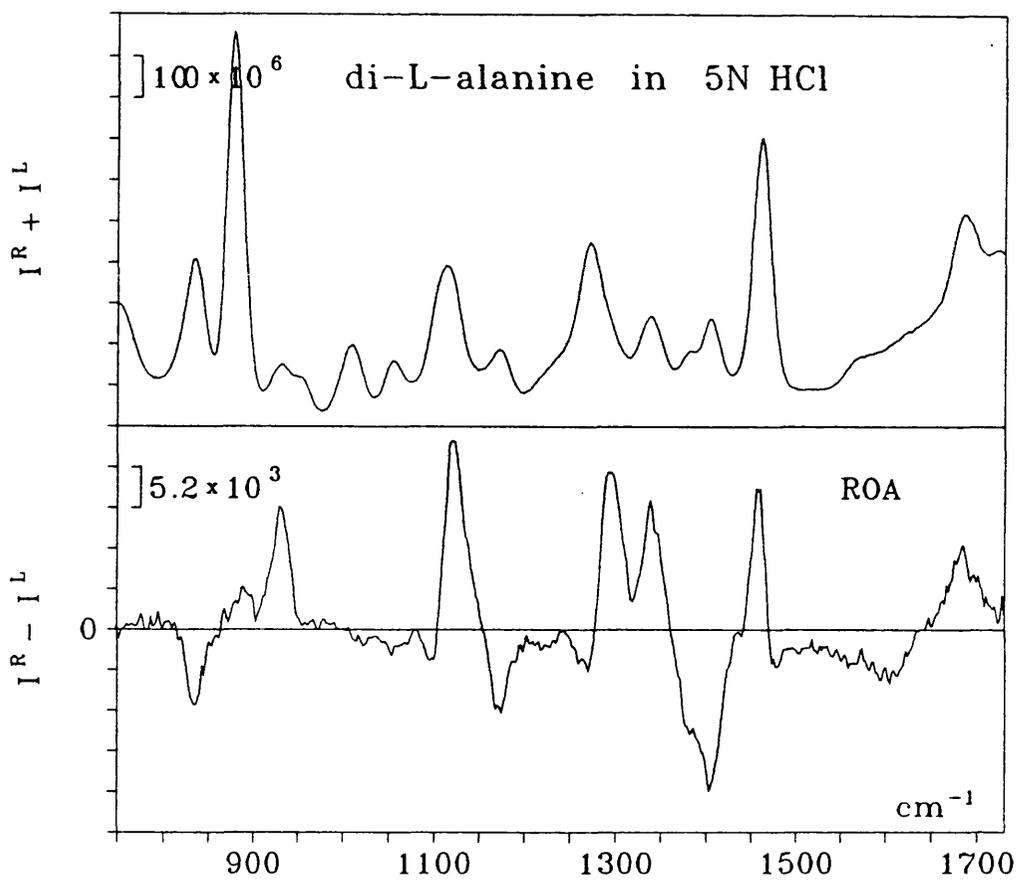


Fig. 4.3 Raman and ROA spectra of di-L-alanine in 5N HCl.

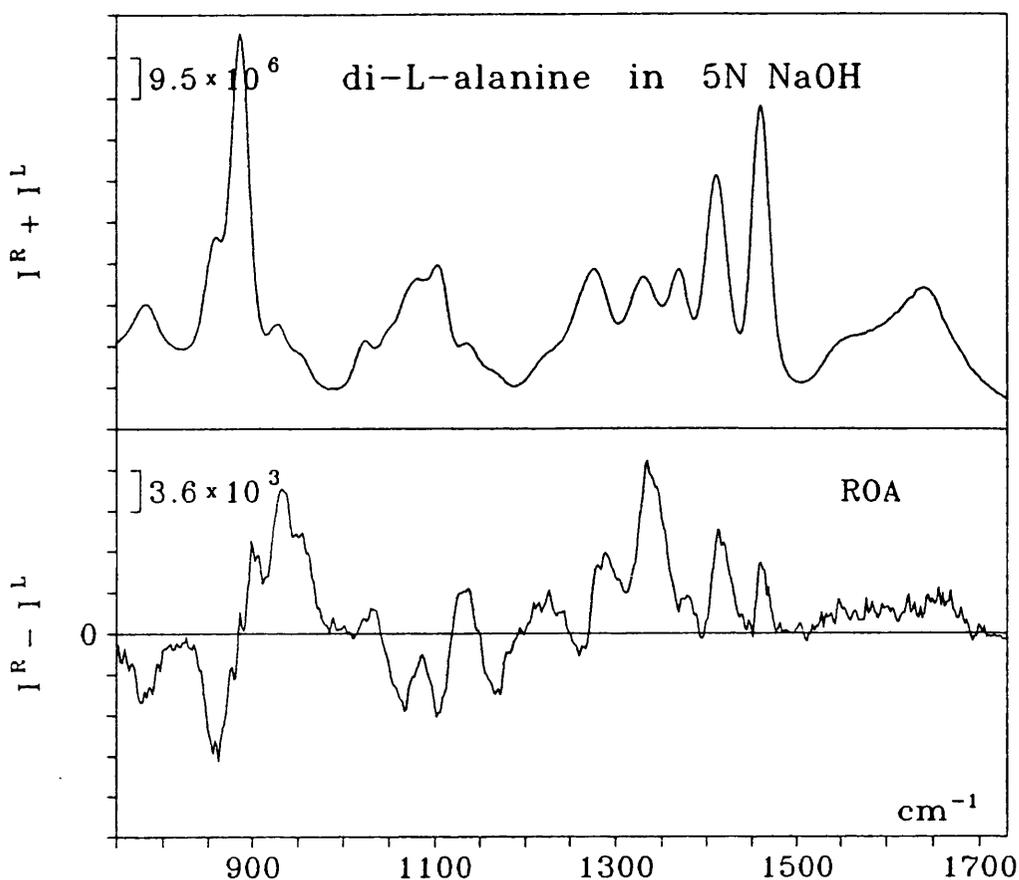


Fig. 4.4 Raman and ROA spectra of di-L-alanine in 5N NaOH.

presented because of the stray light problem which hampered reliable ROA measurement of aqueous solution at lower frequency although good quality ROA spectra can be obtained with organic neat liquids down to  $300\text{cm}^{-1}$ . (We now can measure reliably ROA spectra at lower frequency region with a new holographic notch filter commercially available very recently ).

From Fig. 3.8, one can see that there is very good reflection mirror symmetry of the ROA spectra of the two enantiomers. This ensures the reliability of the ROA data obtained from the ROA instrument. It should

be noted that there are a few rather strongly polarized Raman bands at  $\sim 887\text{cm}^{-1}$ ,  $\sim 1278\text{cm}^{-1}$ , and  $\sim 1465\text{cm}^{-1}$  in the Raman spectrum of the dipeptide. As has been mentioned in the previous chapter 3, strongly polarized Raman bands are often subject to artifact interference. It can be seen from Fig. 3.8 and Figs. 4.2-4.4 that there are no significant artifacts associated with these strongly polarized Raman bands, which indicates that artifacts from strongly polarized Raman bands have been sufficiently suppressed to below the noise level. As a matter of fact, we have used di-L-alanine as an artifact control indicator for ROA measurement of chiral molecules in aqueous solution throughout our instrument development program (*vide supra*) because the ROA signal of di-L-alanine is reasonably strong and its aqueous solution is quite stable for a long period (over one year).

We attempt to divide the entire ROA spectrum of the di-L-alanine into four sub-regions to facilitate the discussion of the ROA bands and to relate them to structure and conformation. ROA feature of polypeptides and proteins measured so far appear chiefly in these regions. The first region is  $1730\text{--}1370\text{ cm}^{-1}$ , which includes the amide I and amide II bands as well as bands associated with terminal groups and side groups. ROA spectra in this region are dominated by the amide I mode; and the amide II mode does not show any significant ROA features. The next region is between  $1370\text{--}1200\text{ cm}^{-1}$  in which are included the so-called amide III bands and which show ample ROA peaks. Vibrational bands in these two regions are associated with relatively local modes. By contrast, vibrational bands in the region below  $1200\text{cm}^{-1}$  usually delocalized and related to the skeleton

vibrational coordinates. It is more difficult to assign bands to corresponding vibrational modes in the lower frequency region. However, it is believed that the  $C_{\alpha}$ -N stretch motion is located between  $1200\text{-}1000\text{cm}^{-1}$  sub-region and the  $C_{\alpha}$ -C stretch and many deformation modes of the chain backbone are in the region below  $1000\text{cm}^{-1}$ .<sup>13</sup> This division is somewhat arbitrary; nevertheless, it may help to better understand the ROA spectra of peptides, polypeptides and proteins.

The Raman and ROA spectra of di-L-alanine can be best compared with its monomer L-alanine. Their parent Raman spectra are overall rather similar. However, the ROA spectrum of di-L-alanine shows striking difference in the region above  $1200\text{cm}^{-1}$  from that of L-alanine although the two ROA spectra expose similarities in the lower frequency region. The interpretation of the ROA spectra of di-L-alanine in the low frequency region can be aided by the elaborate ROA spectra analysis of L-alanine through *ab initio* theoretical calculations.<sup>5</sup> For the Raman and ROA spectra above  $1200\text{cm}^{-1}$ , the discussion is greatly helped by the elegant vibrational Raman spectra analysis of six isotopomers of di-L-alanine via deuteration techniques by Diem et al.<sup>7</sup> To facilitate the discussion of the ROA spectra in the region above  $1200\text{cm}^{-1}$ , we summarized the vibrational analysis of the six isotopomers of the di-L-alanine and their assignment in Table 4-1. In Table 4-2 are compiled the ROA band frequencies and signs of the di-L-alanine in various aqueous solutions and their tentative vibrational assignment.

Table 4-1. The Raman band frequencies of di-L-alanine  
and its deuterated isotopomers between 1750-1200cm<sup>-1</sup>.<sup>a</sup>

CHNHCH <sup>b</sup>	CHNDCH	CDNHCH	CDNDCH	CDNHCD	CDNDCD	assignment
1680	1680			1680		am I C=O str
	1665		1664		1663	am I' C=O str
1570	1592	1570	1590	1570	1584	am II /CO <sub>2</sub> <sup>-</sup> str
	1483		1479		1478	am II'
1460	1458	1457	1460	1458	1458	CH <sub>3</sub> def
	1445		1450		1450	CH <sub>3</sub> def
1407	1407	1407	1408	1406	1406	CO <sub>2</sub> <sup>-</sup> stretch
1372	1368	1369	1371	1370	1370	CH <sub>3</sub> def
	1355					C <sub>N</sub> -H <sup>II</sup> def
1340						amIII <sup>3</sup> (C <sub>α</sub> -H/N-H)
	1346					N-H/C <sub>C</sub> -H def
				1336		amIII (N-H/C-N)
	1329		1330			C <sub>C</sub> -H <sup>II</sup> def
1325						amIII <sup>2</sup> (C <sub>α</sub> -H/N-H)
1280						amIII <sup>1</sup> (C <sub>α</sub> -H/N-H)
		1311				N-H/C <sub>C</sub> -H <sup>II</sup>
	1305					C <sub>N</sub> -H <sup>I</sup> def
	1276	1274	1279			C <sub>C</sub> -H <sup>I</sup> def
1245	1245			1245		NH <sub>3</sub> <sup>+</sup>

a. adapted from the ref. 8.

b. The nomenclature here refers to the di-L-alanine dipeptides isotopomers

CHNHCH, L-ala-L-ala in H<sub>2</sub>O, CHNDCH, L-ala-L-ala in D<sub>2</sub>O;  
 CDNHCH, L-ala-d-L-ala in H<sub>2</sub>O, CDNDCH, L-ala-d-L-ala in D<sub>2</sub>O;  
 CDNHCD, L-ala-d-L-ala-d in H<sub>2</sub>O, CDNDCD, L-ala-d-L-ala-d in D<sub>2</sub>O.

Table 4-2. Raman and ROA band frequencies and signs of di-L-alanine

Raman	ROA(H <sub>2</sub> O)	Raman	ROA(D <sub>2</sub> O)	VCD(D <sub>2</sub> O) <sup>a</sup>	tentative assignment
1684	1688 (+)				am I C=O stretch
		1670	1670 (+)		am I' C=O stretch
	1625 (-)				am I
		1593	1593 (+)	1590 (-)	am I'
1584					am II CO <sub>2</sub> <sup>-</sup> stretch
		1486	1486 (+)		am II'
	1476 (-)				
1465	1460 (+)	1460	1462 (-)		CH <sub>2</sub> , CH <sub>3</sub> def
	1424 (+)				
1410		1410	1410 (-)	1410 (-)	CO <sub>2</sub> <sup>-</sup> stretch
	1402 (-)				
1370	1370 (-)			1370 (-)	CH <sub>3</sub> def, CO <sub>2</sub> <sup>-</sup> stretch
		1360	1360 (-)		C <sub>N</sub> -H def
1340	1340 (+)				N-H/C <sub>C</sub> -H, C <sub>N</sub> -H def
1327		1330	1330 (+)	1330 (-)	C <sub>C</sub> -H def
1280		1280		1280(+)	N-H/C <sub>C</sub> -H, C <sub>N</sub> -H def
	1273 (-)		1273 (+)		C <sub>C</sub> -H def
1175	1175 (-)				C-H bend, CH <sub>3</sub> , NH <sub>3</sub> rock
		1166	1154 (-)		
1108	1100 (-)	1100			C <sub>α</sub> -N, C <sub>α</sub> -CH <sub>3</sub> , CH <sub>3</sub> rock
	1060 (-)	1064	1064 (+)		C <sub>α</sub> -CH <sub>3</sub> , NH <sub>3</sub> rock
1050			1050 (-)		
1012	1015 (-)	1005			NH <sub>3</sub> , CH <sub>3</sub> , rock
958	960 (+)	944			NH <sub>3</sub> , CH <sub>3</sub> rock
930	930(+)	912	915 (+)		C <sub>α</sub> -C(O) C <sub>α</sub> -N stretch
860	862 (-)				C <sub>α</sub> -N bend + COO bend
		825	825 (-)		
782	784 (-)	782			COO wag + COO bend
690	694 (+)				COO bend + COO wag

1730-1370 $\text{cm}^{-1}$ . We start to discuss the ROA spectrum together with the Raman spectrum of di-L-alanine from the high frequency side. Vibrational bands of di-L-alanine in this region can be readily assigned without much controversy since they are associated with local vibrational coordinates and are basically group frequencies. There are five Raman bands in this region:  $\sim 1684\text{cm}^{-1}$ ,  $\sim 1584\text{cm}^{-1}$ ,  $\sim 1465\text{cm}^{-1}$ ,  $\sim 1410\text{cm}^{-1}$  and  $\sim 1370\text{cm}^{-1}$ . The band at  $\sim 1684\text{cm}^{-1}$  is due to the so-called amide I band which is contributed dominantly by the C=O stretch coordinate of the peptide bond. It shows a negative and positive ROA couplet, negative at lower frequency  $\sim 1625\text{cm}^{-1}$  (-) and positive at higher  $\sim 1660\text{cm}^{-1}$ (+). No significant ROA feature is observed at the amide I band position in the corresponding monomer L-alanine.<sup>5</sup> This suggests that the ROA couplet is generated by the C=O stretch coordinate of the peptide linkage through coupling with other vibrational modes. Neither of the two molecules show any significant VCD feature at the amide I' band in  $\text{D}_2\text{O}$  within the limit of the current VCD instrument sensitivity.<sup>8</sup>

It is intriguing to note the negative-positive sign pattern of this ROA couplet since all proteins measured so far exhibit a large ROA negative-positive couplet associated with the amide I band. The couplet is also present in the ROA spectra of alanine oligomers. However, denatured protein and the polyamino acids poly-L-lysine and poly-L-glutamic acid in ionized form expose a positive ROA peak at the amide I band position (*vide infra*). It should be recalled that di-L-alanine exists in an extended conformation in aqueous solution.<sup>4</sup>

The next Raman band at  $\sim 1570\text{cm}^{-1}$  is very weak but strong in

infrared absorption and is assigned straightforwardly to the amide II mode arising from the N-H in plane bend and C-N stretch. No significant ROA feature is observed at the amide II band region. Upon N-deuteration, this band shifts to lower frequency at  $\sim 1480\text{cm}^{-1}$  and shows very strong Raman intensity and gives rise to a positive ROA peak. VCD spectra of L-alanine and di-L-alanine do not show any features at the amide II' band.<sup>8</sup>

The three remaining Raman bands at  $\sim 1465\text{cm}^{-1}$ ,  $\sim 1410\text{cm}^{-1}$  and  $\sim 1370\text{cm}^{-1}$  in this region are due to the antisymmetric methyl deformation, the  $\text{CO}_2^-$  symmetric stretch motion and the symmetric methyl deformation plus  $\text{CO}_2^-$  stretch vibration respectively. These Raman bands remain nearly invariant between the monomer and the dimer and upon deuteration at N-H and C-H in the isotopomers.<sup>7</sup> There is a positive and negative ROA couplet associated with the Raman band at  $\sim 1465\text{cm}^{-1}$  in di-L-alanine. By contrast, the ROA spectrum of L-alanine presents an ROA couplet at this position but with an opposite sign pattern. This couplet in L-alanine is suggested to be generated in the two split antisymmetric methyl deformation through coupling with other modes.<sup>5</sup> The change of sign pattern of the couplet from the monomer to the dimer probably indicates that the methyl groups in di-L-alanine are in a different local environment from that of L-alanine.

There is a negative-positive ROA couplet related to the Raman band at  $\sim 1410\text{cm}^{-1}$ . A tiny negative feature at this frequency is shown in the VCD spectra of di-L-alanine in water.<sup>8</sup> The Raman band at this frequency has been assigned to the  $\text{CO}_2^-$  symmetric stretch and the symmetric methyl deformation in L-alanine.<sup>5</sup> However, the Raman and ROA spectra of

di-L-alanine in 5N HCl shows that the Raman band intensity is now greatly reduced and the ROA couplet becomes a large negative band (Fig. 4.3). These data support the assignment of the  $\sim 1410\text{cm}^{-1}$  to the vibration of the terminal group  $\text{CO}_2^-$  and exclude the contribution from the methyl deformation.

The Raman band at  $\sim 1370\text{cm}^{-1}$  has been assigned to the symmetric methyl deformation plus  $\text{CO}_2^-$  symmetric stretch<sup>7</sup> which occurs at very similar frequency in L-alanine. It shows a negative ROA feature at  $\sim 1370\text{cm}^{-1}(-)$ . Upon N-deuteration, the negative ROA feature disappears, which indicates some involvement of the N-H group in this band. VCD spectra of di-L-alanine shows a small negative feature at  $\sim 1375\text{cm}^{-1}(-)$ .<sup>8</sup>

**1370-1200 $\text{cm}^{-1}$  region.** We now proceed to discuss the Raman and ROA spectra in the extended amide III band region. ROA features in this region probably contain the most significant information associated with the peptide chain conformation since all proteins measured so far exhibit characteristic ROA features in this region. To understand the origin of ROA features of di-L-alanine in this region has an immediate implication in proteins.

The ROA spectrum of di-L-alanine in this region is considerably different from that of L-alanine in contrast to the similarity in the lower frequency region (*vide infra*). It exhibits a huge positive peak at  $\sim 1340\text{cm}^{-1}(+)$  and a negative band at  $\sim 1273\text{cm}^{-1}(-)$ . This typical ROA pattern also presents in the ROA spectra of the oligomers tri-L-alanine and tetra-L-alanine (*vide infra*). Contrary to the di-L-alanine, the L-alanine

only exposes two small positive bands at  $1351\text{cm}^{-1}(+)$  and  $1301\text{cm}^{-1}(+)$  respectively.<sup>5</sup>

Large VCD signals have also been recorded for di-L-alanine in water in this region.<sup>8</sup> The most prominent VCD feature is a strong positive peak at  $\sim 1280\text{cm}^{-1}(+)$ . There is a distinct negative band at  $\sim 1325\text{cm}^{-1}(-)$  and a small negative signal at  $\sim 1375\text{cm}^{-1}(-)$ .

The vibrational modes associated with these ROA features are quite complex. There are four apparent Raman bands presented in this region. The band at  $\sim 1330\text{cm}^{-1}$  is overlapped with a shoulder band at  $\sim 1340\text{cm}^{-1}$ . The band at  $\sim 1280\text{cm}^{-1}$  had been previously assigned to the so-called amide III mode in peptides, which was proposed to arise mainly from the N-H in plane deformation and C-N stretch. These Raman bands have been re-assigned in the studies by Diem et al.<sup>7</sup> (Table 4-1 ).

The broad strong Raman band at  $\sim 1280\text{cm}^{-1}$  is believed to be a superposition of four components at  $\sim 1245\text{cm}^{-1}$ ,  $\sim 1266\text{cm}^{-1}$ ,  $\sim 1280\text{cm}^{-1}$  and  $\sim 1302\text{cm}^{-1}$  by band deconvolution.<sup>7</sup> The other overlapped band  $\sim 1330\text{cm}^{-1}$  and the shoulder band is separated into two bands at  $\sim 1325\text{cm}^{-1}$  and  $1340\text{cm}^{-1}$  respectively. The band at  $\sim 1245\text{cm}^{-1}$  is thus assigned, in analogy to alanine, as a  $\text{NH}_3^+$  rocking mode, which is observed in all di-L-alanine isotopomers in water. The two peaks at  $\sim 1266\text{cm}^{-1}$  and  $\sim 1302\text{cm}^{-1}$  are consequently designated to the two uncoupled methine deformations respectively since similar Raman bands appear at  $\sim 1276\text{cm}^{-1}$  and  $\sim 1305\text{cm}^{-1}$  in the N-deuterated di-L-alanine in  $\text{D}_2\text{O}$  in which the methine deformation is uncoupled with the deuteration at the N-H group.

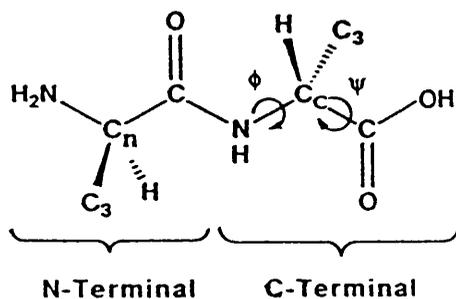


Fig. 4.5 An illustration of the  $\varphi$  and  $\psi$  angles of the peptide linkage and the C, N terminal groups in di-L-alanine.

The rest of three Raman bands  $\sim 1340\text{cm}^{-1}$ ,  $\sim 1325\text{cm}^{-1}$  and  $\sim 1280\text{m}^{-1}$  are the most interesting and are all associated with the so-called amide III mode. They are designated by Diem et al.<sup>7</sup> as amide III<sup>3</sup>, amide III<sup>2</sup>, and amide III<sup>1</sup> respectively and are all highly coupled modes arising from the N-H in-plane bend, the  $\text{C}_C\text{-H}^{\text{II}}$  deformation and the  $\text{C}_N\text{-H}^{\text{II}}$  deformation coordinates. Among the three amide III bands, the  $\sim 1280\text{cm}^{-1}$  band is the most delocalized mode and has contributions from N-H in plane deformation and C-N stretch plus the two methine deformations. It represents the most highly mixed of all the amide III modes and disappears upon deuteration of methine or amide hydrogen. The two observed overlapped Raman bands at  $\sim 1340\text{cm}^{-1}$  and  $\sim 1325\text{cm}^{-1}$  are assigned to the two other linear combination of the three coupling coordinates, namely the N-H deformation and the two C-H deformation respectively. Normal mode calculation<sup>9</sup> reveals that the former (amide III<sup>3</sup>)

has a large contribution from N-H in plane deformation, coupled mostly to the  $C_N-H^{II}$  vibration and the latter (amide III<sup>2</sup>) consists chiefly of the  $C_C-H^{II}$  deformation, coupled with the  $C_N-H^{II}$  but with a negligible contribution from the N-H deformation.

With the detailed assignment of these Raman bands in this region, the origin of the ROA features are readily correlated to their corresponding vibrational modes. Thus, the huge positive ROA band at  $\sim 1340\text{cm}^{-1}(+)$  is generated by the coupling of the N-H in plane deformation and the  $C_N-H^{II}$  deformation and the negative ROA feature at  $\sim 1273\text{cm}^{-1}(-)$  is attributed exclusively to the  $C_C-H^I$  deformation. The amide III<sup>2</sup> mode probably contributes a small portion of ROA intensity to the lower frequency side of the huge positive ROA band centered at  $\sim 1340\text{cm}^{-1}(+)$  which is obviously asymmetric in shape ( see Fig. 3.8 ). Likewise, prominent VCD features at  $\sim 1280\text{cm}^{-1}$ ,  $\sim 1325\text{cm}^{-1}$  are also related with the amide III<sup>1</sup> and amide III<sup>2</sup> modes respectively.

The ROA spectrum of di-L-alanine shows dramatic changes upon N-deuteration (Fig. 4.2). The huge positive ROA peak at  $1340\text{cm}^{-1}(+)$  has now split into two ROA bands, one positive peak at  $1330\text{cm}^{-1}(+)$  and one negative band at  $\sim 1360\text{cm}^{-1}(-)$ . The negative ROA feature at  $\sim 1273\text{cm}^{-1}(-)$  has changed its sign from negative to positive. These deuterated ROA data confirm that the vibrational modes responsible for the Raman band at  $\sim 1340\text{cm}^{-1}$  is associated with the amide N-H group and provide further evidence for the coupling between the N-H and C-H deformations. ROA spectra obtained in 5N HCl and 5N NaOH reveal that, apart from the dominant contribution from the amide III modes, the terminal group  $\text{CO}_2^-$

also contributes to the Raman band at  $\sim 1290\text{cm}^{-1}$  because protonation of the di-L-alanine results in a huge positive ROA peak at  $\sim 1290\text{cm}^{-1(+)}$  in 5N HCl solution (Fig. 4.3). This is in agreement with the band assignment from  $^{18}\text{O}$  isotopic substitution study of L-alanine that the IR vibrational band at  $\sim 1300\text{cm}^{-1}$  has contribution from the  $\text{CO}_2^-$  symmetric stretch.<sup>12</sup>

This new understanding of the so-called amide III mode explains the mechanism of the great geometric sensitivity of the amide III band. It is the coupling between the N-H group at the peptide linkage and the  $\text{C}_\alpha\text{-H}$  group at the asymmetric carbon atom that give rise to the conformational sensitivity. This interpretation agrees well with the results from previous normal mode calculation of polypeptide chain by Krimm et al.<sup>10</sup> that there is considerable mixing between N-H in plane bend and the  $\text{C}_\alpha\text{-H}$  deformation, and the amide III band is sensitive to the dihedral angle  $\varphi$  rather than  $\psi$ . The large ROA signal at  $\sim 1340\text{cm}^{-1(+)}$  generated by the coupling mechanism between N-H and  $\text{C}_\alpha\text{-H}$  groups further supports the vibrational analysis and probably offers a more sensitive probe to study peptide chain backbone conformation. Hopefully this ROA feature can be used to probe the change of the  $\varphi$  angle, which plays a crucial role in determining the peptide chain backbone conformation, if the correlation between the ROA intensity of the amide III<sup>3</sup> band and the  $\varphi$  angle can be established.

It is intriguing to compare the ROA features of the amide III modes with corresponding VCD bands. The large positive ROA band at  $\sim 1340\text{cm}^{-1(+)}$  contrasts with a vanishing VCD at the same frequency in  $\text{H}_2\text{O}$ . Upon N-deuteration in  $\text{D}_2\text{O}$ , the uncoupled  $\text{C}_\alpha\text{-H}$  deformation gives

rise to a large negative VCD band at  $\sim 1330\text{cm}^{-1}(-)$ <sup>8</sup> and a large positive ROA peak at the same frequency and a negative peak at  $\sim 1360\text{cm}^{-1}(-)$ . The large positive VCD band at  $\sim 1280\text{cm}^{-1}(+)$  is considerably reduced in intensity and the negative ROA feature at this frequency changes its sign to positive when coupling between the N-H and the  $\text{C}_\alpha\text{-H}$  deformation occurs in  $\text{D}_2\text{O}$ . These observations indicate that the mechanism of generating ROA is quite different from VCD and further emphasises that these two techniques offer highly complementary information on molecular conformation. It is gratifying that the simple two group model of vibrational optical activity predicts that coupled deformation vibrations in the two group structure produce ROA signal but not VCD, at least within the approximation that there is no change of dipole moment during the deformation motion of the groups.<sup>11</sup>

**1200-1000 $\text{cm}^{-1}$  region.** The major ROA features of di-L-alanine in this region consist of two couplets. One couplet is negative at  $\sim 1100\text{cm}^{-1}$  and positive at  $\sim 1121\text{cm}^{-1}$ , which coalesce with the positive peak of the adjacent couplet. The negative part of the second couplet is at  $\sim 1175\text{cm}^{-1}$ . This ROA pattern is also present in tri-L-alanine and tetra-L-alanine.

Conventional Raman bands of peptides in this region have been correlated to skeleton vibrational modes including  $\text{C}_\alpha\text{-N}$  stretch and  $\text{C}_\alpha\text{-C}$  stretch coordinates.<sup>13</sup> The detailed analysis of L-alanine has revealed that ROA features in this region involve  $\text{NH}_3^+$  and  $\text{CH}_3$  rock and the C-N and  $\text{C}_\alpha\text{-CH}_3$  stretch.<sup>5</sup>

Upon N-deuteration, the ROA pattern in this region becomes two well

resolved couplets. ROA spectra of di-L-alanine in HCl and NaOH solution show that the ROA band at  $\sim 1120\text{cm}^{-1}$  is pH dependent. This indicates that the band is related to the C-terminal group  $\text{CO}_2^-$  vibration, which is in good agreement with the conclusion inferred from infrared spectra of  $^{18}\text{O}$  substituted L-alanine.<sup>12</sup>

It should be noted that the ROA pattern in this region is more similar to the alanyl oligomers than the monomer L-alanine. Moreover, polypeptides and proteins generally exhibit a negative-positive feature at  $1050\text{--}1150\text{cm}^{-1}$ . This indicates that apart from the contribution from the side groups  $\text{CH}_3$  and  $\text{NH}_3^+$ , the peptide linkage plays a significant role in generating ROA features in this region.

**1000-750 $\text{cm}^{-1}$  region.** The di-L-alanine shows a very similar ROA pattern to that of L-alanine in the region between  $1000\text{--}750\text{cm}^{-1}$ . It exhibits a small negative band at  $\sim 784\text{cm}^{-1}(-)$ , a huge negative peak at  $\sim 862\text{cm}^{-1}(-)$  and a positive peak at  $\sim 930\text{cm}^{-1}(+)$ . All the three band frequencies are slightly higher than the corresponding ones in L-alanine in which they occur at  $\sim 775\text{cm}^{-1}(-)$ ,  $\sim 850\text{cm}^{-1}(-)$  and  $\sim 922\text{cm}^{-1}(+)$  respectively. These ROA features presumably arise from similar vibrational modes involving the C-C stretch and the symmetric COO bend and  $\text{C}_\alpha\text{-N}$  stretch coordinates as in L-alanine.<sup>5</sup> The  $\sim 775\text{cm}^{-1}$  band in L-alanine has been assigned to the COO wagging plus symmetric COO bend. The ROA band at  $\sim 862\text{cm}^{-1}(-)$  in water shifts to  $\sim 825\text{cm}^{-1}(-)$  and  $\sim 832\text{cm}^{-1}(-)$  upon N-deuteration and acidification respectively. This observation further confirm the assignment of this band to the  $\text{C}_\alpha\text{-N}$  bend and  $\text{CO}_2^-$  group.

It is quite interesting to note that the positive ROA feature at  $\sim$

930cm<sup>-1</sup> has a relatively large ROA  $\Delta$  value (  $\sim 2 \times 10^{-3}$  ) and also appears in the ROA spectra of the tri-L-alanine and tetra-L-alanine oligomers. The Raman band at this frequency has been suggested to be a helical structure marker in polypeptides.<sup>13</sup> Proteins with high fraction of  $\alpha$ -helix components usually show a positive ROA feature at the frequency region 900-950cm<sup>-1</sup> but proteins containing high  $\beta$ -sheet do not exhibit any significant ROA in this area. It has been long recognized that L-alanine and alanyl oligomers promote and stabilize  $\alpha$ -helix chain conformation.

#### 4.3. Raman optical activity of tri- and tetra-L-alanine

It is natural to extend the ROA measurement to the alanine oligomer tri-L-alanine and tetra-L-alanine to investigate chain length dependence effect. Tri-L-alanine has been recognized to be able to inhibit the esterolytic activity of elastase<sup>14,15</sup> and tetra-L-alanine has been used as a model peptide for the calculation of characteristic vibration mode of the reverse turn structure.<sup>16</sup>

The ROA spectra of tri-L-alanine in water is shown in Fig. 4.6. The ROA spectra of tetra-L-alanine is obtained in diluted aqueous acid solution (0.1M HCL) and is shown in Fig. 4.7 (tetra-L-alanine does not dissolve well in water at neutral pH but increases solubility at low pH ).

The overall ROA spectra of tri- and tetra-L-alanine in water are very similar to that of di-L-alanine. The most interesting ROA features of the alanine oligomers are in the extended amide III region where di-L-alanine shows a large positive ROA peak at  $\sim 1340\text{cm}^{-1}(+)$  which shifted to lower

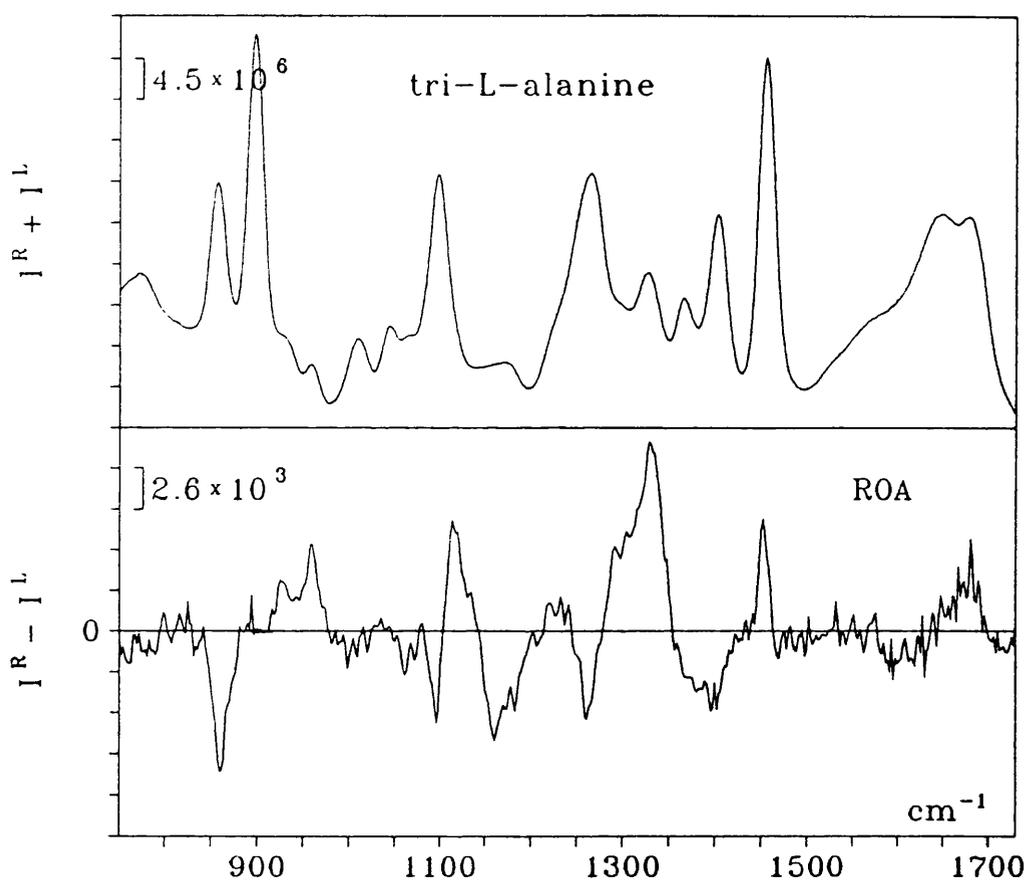


Fig. 4.6 Raman and ROA spectra of tri-L-alanine in water.

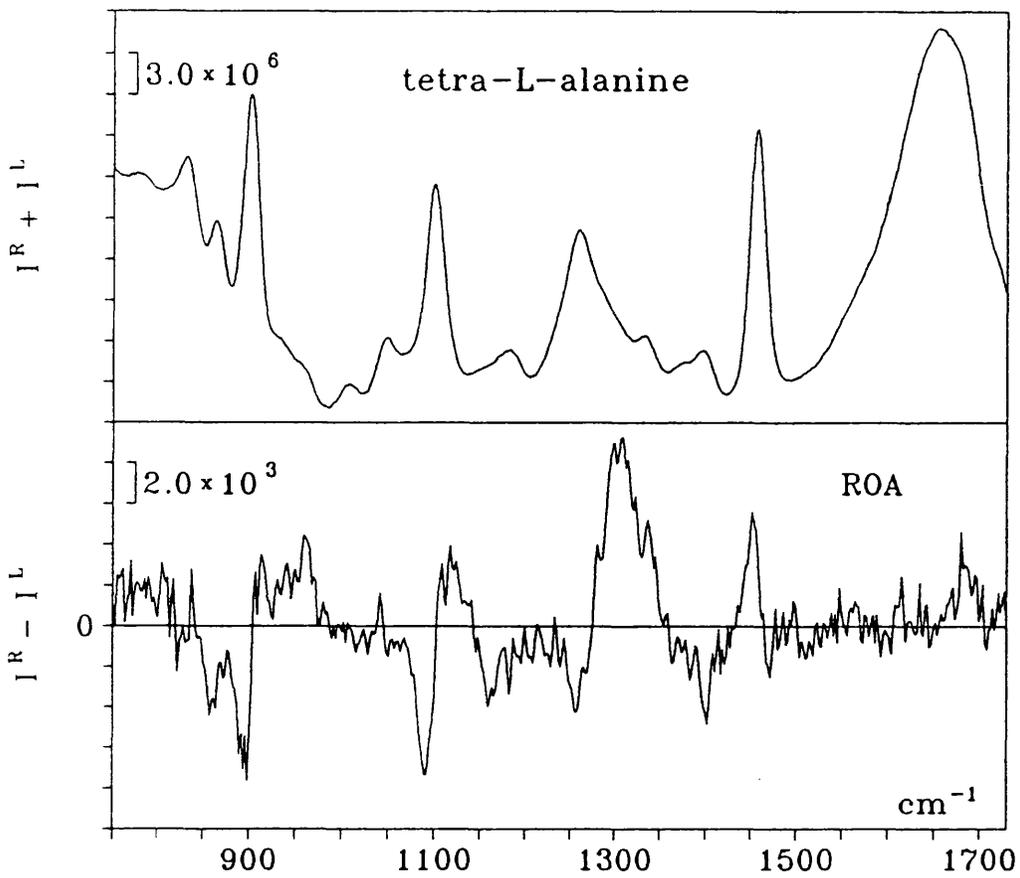


Fig. 4.7 Raman and ROA spectra of tetra-L-alanine in 0.1M HCl.

Table 4-3. Raman and ROA band frequencies  
and signs of tri-L-alanine and tetra-L-alanine

tri-L-alanine		tetra-L-alanine		tentative assignment
Raman	ROA	Raman	ROA	
1682	1682 (+)		1684 (+)	am I C=O stretch
1650		1658		am I C=O stretch
	1620 (-)			
1456	1470 (-)	1457		
	1452 (+)		1452 (+)	CH <sub>2</sub> , CH <sub>3</sub> def
1401	1398 (-)		1402 (-)	
1367		1397		CH <sub>3</sub> def CO <sub>2</sub> <sup>-</sup> stretch
1327	1330 (+)	1334	1337 (+)	am III, C <sub>α</sub> -H/N-H def
	1300 (+)		1305 (+)	am III, C <sub>α</sub> -H/N-H def
1267	1261 (-)	1262	1257 (-)	C <sub>C</sub> -H def
1227	1226 (+)			
1174	1161 (-)	1184	1161 (-)	C-H bend, CH <sub>3</sub> , NH <sub>3</sub> rock
	1116 (+)		1118 (+)	C <sub>α</sub> -N, C <sub>α</sub> -CH <sub>3</sub> stretch
1097	1096 (-)	1099	1090 (-)	
1064				
1045		1049		
1010		1010		NH <sub>3</sub> , CH <sub>3</sub> rock
960	960 (+)		961 (+)	NH <sub>3</sub> , CH <sub>3</sub> rock
929				C <sub>α</sub> -C, C <sub>α</sub> -N stretch
897		901		
856	860 (-)	862	857 (-)	
773		782		
		752		

frequency at  $1330\text{cm}^{-1}(+)$  together with a positive shoulder at  $\sim 1300\text{cm}^{-1}(+)$  in tri-L-alanine. The frequency shift of the amide III<sup>3</sup> mode with increase of the peptide chain length indicates that the mode becomes delocalized along the peptide chain since the amide III<sup>3</sup> involves the coupling of the N-H and C<sub>α</sub>-H deformation separated by two bonds.

The negative ROA feature at  $\sim 1273\text{cm}^{-1} (-)$  in the di-L-alanine remains unchanged in tri-L-alanine and tetra-L-alanine. This further confirms the assignment of this band to local C<sub>α</sub>-H deformation plus some contribution from N-H deformation and indicates it is relatively localized.

It is surprising that the ROA spectrum of tri-L-alanine in D<sub>2</sub>O is very similar to that of di-L-alanine (Fig. 4.2) except a small difference of relative magnitude in some bands. The two ROA features at  $1330\text{cm}^{-1}(+)$  and  $\sim 1360\text{cm}^{-1}(-)$  in tri-L-alanine are almost identical to the corresponding di-L-alanine which have been related to the uncoupled C<sub>α</sub>-H deformation. The deuterated ROA data confirm that the vibration modes of these two ROA bands are relatively localized: the proton in the N-H group in water is crucial in the coupling of the C<sub>α</sub>-H deformation along the peptide chain because the replacement of the proton in the N-H group by a deuterium in D<sub>2</sub>O now apparently break up the coupling between the C<sub>α</sub>-H deformations. However, some weaker coupling may still remain between the C<sub>α</sub>-H deformations because the small positive ROA feature at  $\sim 1310\text{cm}^{-1}(+)$  in di-L-alanine in D<sub>2</sub>O increases its intensity in the tri- and tetra-L-alanines.

#### 4.4 Raman optical activity of L-Pro-L-Leu-Gly-amide

Reverse turn structure is one of the important secondary structure components of proteins together with  $\alpha$ -helix,  $\beta$ -sheet and random coil. It is particularly significant in globular proteins in which about one-third of amino acid residues occur as reverse turns. They play a vital role in the polypeptide folding processes leading to the native conformation of proteins and make an otherwise extended polypeptide chain compact. It has been observed that a substantial portion of surface residues, where most of enzymatic reaction taken place, are in reverse turns. It is important to investigate the model reverse turn structure to help the interpretation of protein ROA.

The tripeptide L-pro-L-leu-gly-amide is a useful model peptide for  $\beta$ -turn structure.<sup>16</sup> It is from the C-terminal of oxytocin and has the ability of inhibiting the release of melanocyte-stimulating hormone. The molecule has an intramolecular hydrogen bond between the carbonyl of the proline residue and the trans carboxamide proton of glyciamide and thus forms a rigid ten membered loop structure called  $\beta$ -turn II. The  $\beta$ -turn II structure of this peptide has been confirmed by X-ray diffraction<sup>17</sup> in crystal and NMR in DMSO solution.<sup>18</sup> Raman spectra of the molecule in crystal and aqueous solution have been previously reported.<sup>19, 20</sup> Normal mode calculations have been carried out on it to establish the frequency-structure correlation for reverse turns. A schematic view of the molecular structure is shown in Fig. 4.8.

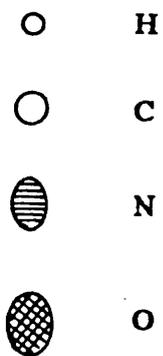
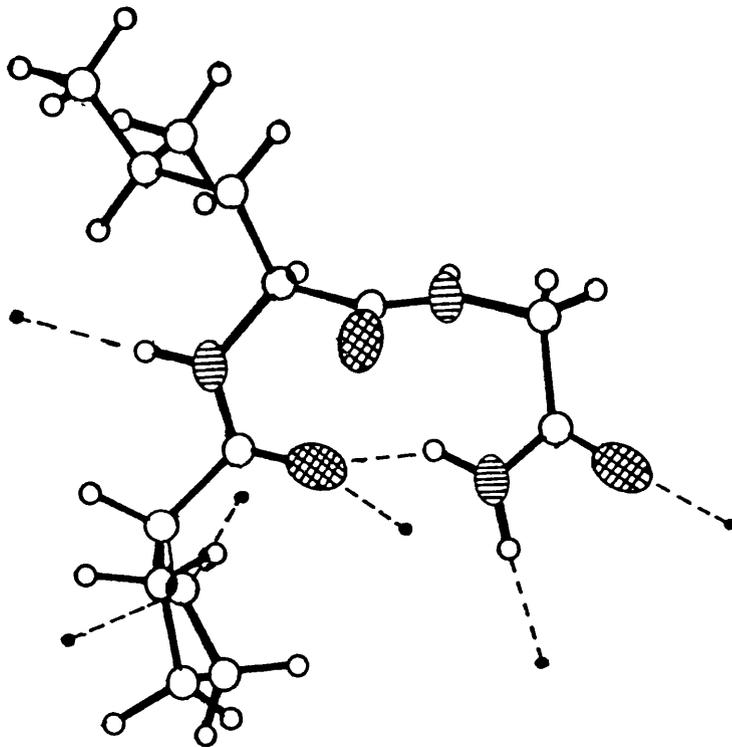


Fig. 4.8 A schematic view of the backbone conformation of the tripeptide L-pro-L-leu-gly-amide.

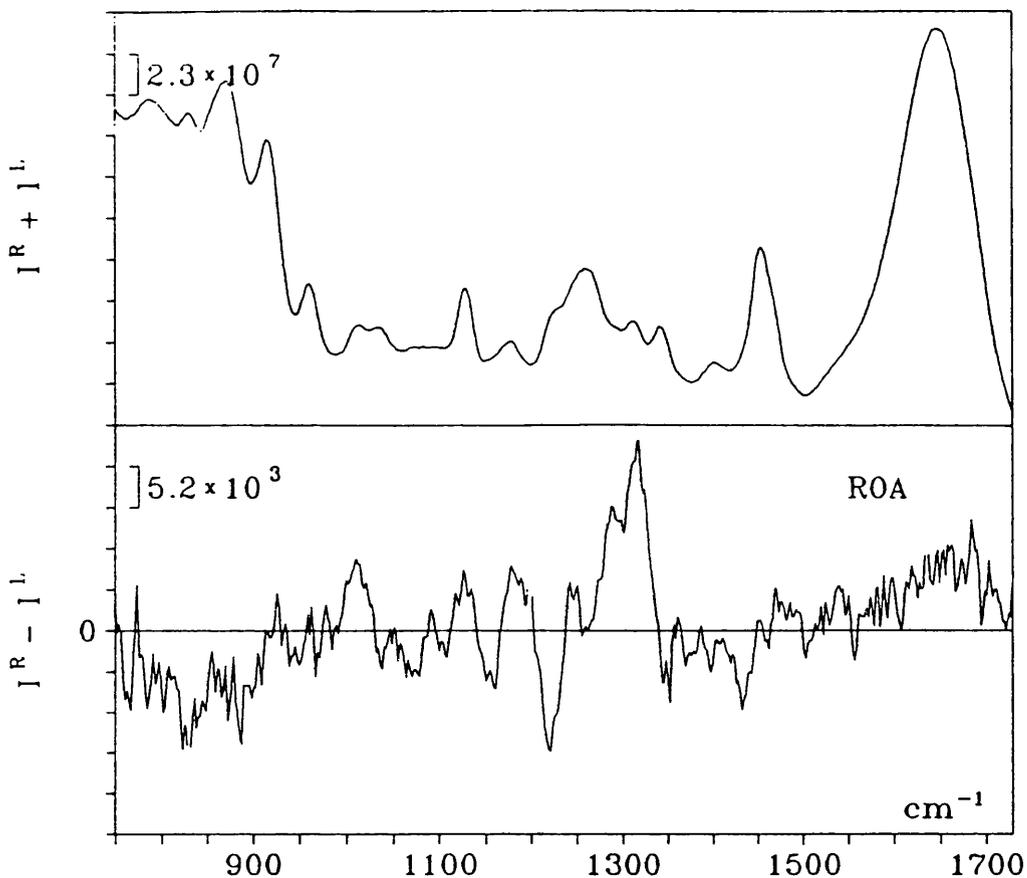


Fig. 4.9 Raman and ROA spectra of L-Pro-L-Leu-Gly-amide in water.

The Raman and ROA spectra of the tripeptide L-pro-L-leu-gly-amide was recorded with a sample concentration of  $\sim 0.4\text{M}$  in water (Fig. 4.9). A total acquisition time of 36 hours was required since the Raman spectrum is rather weak. However, major ROA features are clearly discernible though the ROA spectrum quality is not as good as other peptides and proteins. For the convenience of the discussion of the ROA spectra, the Raman and ROA band frequencies and their tentative assignment are summarized in Table 4-4.

The most prominent ROA features are at the extended amide III

Table 4-4. Raman and ROA band frequencies of L-Pro-L-Leu-Gly-amide

Raman	ROA	tentative assignment <sup>a</sup>
1647		C=O stretch and C-C-N def
1454	1431 (-)	CH <sub>2</sub> and CH <sub>3</sub> def
1405		
1342	1350 (-)	CH <sub>2</sub> and CNH def, NH bend
	1330 (+)	C-N stretch, N-H, C-H def
1312		
	1290 (+)	
1261	1257 (-)	CH <sub>2</sub> def, NH in-plane bend
1226	1220 (-)	C-C and C-N stretch
1177	1180 (+)	C-C and C-N stretch
1129	1126 (+)	NH <sub>2</sub> def, C-N and C-C stretch
	1074 (-)	
1034		C-N stretch, CH <sub>3</sub> and CH <sub>2</sub> def
1015	1012 (+)	CH <sub>3</sub> def
961		C-C and C-N stretch, CH <sub>2</sub> def
873	880 (-)	C-C and C-N stretch, CNH def
829	824 (-)	CH <sub>2</sub> def, C-C stretch

a. adapted from reference 18.

region. However, a broad weak positive ROA feature can be seen in the amide I region which is related to the C=O stretch coordinate of the peptide linkage. A negative band at  $\sim 1430\text{cm}^{-1}(-)$  could probably be associated with the methylene group.

In the extended amide III region, there are two strong positive bands

at  $\sim 1330\text{cm}^{-1}(+)$  and  $\sim 1290\text{cm}^{-1}(+)$  and a large negative band at  $\sim 1220\text{cm}^{-1}(-)$ . Since normal mode calculations predict a vibrational mode at  $\sim 1340\text{cm}^{-1}$  which involves N-H in plane deformation plus  $\text{C}_\alpha\text{-H}$  deformation of proline and leucine residues.<sup>20</sup> We suggest that the positive peak at  $\sim 1330\text{cm}^{-1}(+)$  is attributed to the N-H in-plane deformation and  $\text{C}_\alpha\text{-H}$  deformation. The Raman spectrum clearly shows that this band vanishes upon N-deuteration,<sup>21</sup> which confirms that the band has significant contributions from the N-H group in proline and leucine residues.

The most intriguing ROA feature is a small negative band at  $\sim 1350\text{cm}^{-1}(-)$  in the extended amide III region. The vibrational origin of this band could be associated with the  $\text{CH}_2$  deformation and N-H related mode in accordance with a vibrational normal mode analysis.<sup>18</sup> The negative sign of this ROA band is particularly appealing because proteins containing a high fraction of  $\beta$ -sheet and  $\beta$ -turn conformation show a similar negative ROA peak at this frequency but none appears in the other proteins (*vide infra*). Of course, it is premature at the moment to correlate this negative ROA feature to the characteristic  $\beta$ -turn structure. Nevertheless, it may provide a clue to reveal the characteristic ROA band of reverse turn structure in peptides and proteins.

The issue of characteristic Raman bands for  $\beta$ -turn structure in peptides and proteins have been much debated.<sup>10, 21-24</sup> Krimm et al. have proposed that the characteristic amide III band for reverse turn structure appears above  $1300\text{cm}^{-1}$ , which is apparently beyond the conventional amide III region, based on the vibrational normal mode calculation of a number

of model  $\beta$ -turn peptides.<sup>10</sup> They further suggested that many Raman bands in the region  $\sim 1300\text{cm}^{-1}$  could be assigned to reverse turn components. On the contrary, Vergoten et al.<sup>22-24</sup> argued that the characteristic  $\beta$ -turn amide III band frequency is within  $1250\text{--}1280\text{cm}^{-1}$ , which coincide with the characteristic amide III band frequency of the  $\alpha$ -helix.

Raman spectrum of the tripeptide L-pro-L-leu-Gly-amide in deuterated form has been reported by Fox et al.<sup>21</sup> Their interpretation of the  $\beta$ -turn amide III band is that they overlap with amide III bands of other secondary structure in the region  $\sim 1240\text{cm}^{-1}$  and claim that no amide III Raman bands were observed above  $1300\text{cm}^{-1}$ . However, the Raman spectrum clearly show that there are three Raman bands between  $1300\text{cm}^{-1}$  and  $1400\text{cm}^{-1}$  region and they disappear upon N-deuteration, which supports the assignment of these Raman bands to N-H related vibrational modes.

Our ROA data show that strong ROA features are present between  $1300\text{cm}^{-1}$  and  $1360\text{cm}^{-1}$  and they are related to N-H groups, which support the proposal of Krimm et al. that characteristic amide III band of reverse turns are above  $1300\text{cm}^{-1}$ . Because ROA is a local effect and is very sensitive to the rigidity of the molecular frame and the reverse turn structure of peptides are rather rigid.

It is expected that ROA measurement on this  $\beta$ -turn model peptide in  $\text{D}_2\text{O}$  would give rise to more information on the characteristic ROA band of  $\beta$ -turn structure. Unfortunately, the Raman spectrum is rather weak and we could not measure it in  $\text{D}_2\text{O}$ . It is hoped that ROA spectra of this model peptide, together with other  $\beta$ -turn model peptides, would be investigated when the instrument sensitivity is further improved.

#### 4.5. Raman optical activity of poly-L-lysine and poly-L-glutamic acids

Polyamino acids are synthetic polymers composed of amino acid residues joined by peptide bonds. They are ideal model polypeptides for protein structure and have served as test compounds for assessment of the usefulness and reliability of many techniques for protein study.<sup>25</sup> Among the many polyamino acids available, poly-L-lysine (PLL) and poly-L-glutamic acid (PLG) are particularly favoured by spectroscopists (ORD, ECD, Raman, Infrared, VCD)<sup>25-28</sup> because they can form a single chain conformation under certain experimental conditions and change from one conformation to the other upon change of the environment such as pH and temperature.

The ORD and ECD spectra of PLL and PLG have been extensively investigated from both experimental and theoretical aspects. It is the elaborate analysis of these two polyamino acids by ORD and ECD that established the standard method to estimate statistically the fraction of the secondary structure components in proteins in solution phase.<sup>26</sup> The ECD spectra of PLL has been embodied into contemporary biochemistry texts as a classical example to illustrate the usefulness of ECD spectroscopy in discriminating the three basic components of protein secondary structure, namely,  $\alpha$ -helix,  $\beta$ -sheet and random coil.

Similar to ECD, conventional Raman spectroscopy<sup>27</sup> has used PLL and PLG as model polypeptide to justify the usefulness of Raman spectra in protein structure determination. Their characteristic Raman band frequencies of the peptide linkage are well correlated with the polypeptide

chain backbone conformation through intrachain hydrogen bonding (in  $\alpha$ -helix) and interchain hydrogen bonding (in  $\beta$ -sheet). Based on the three-component model of protein secondary structure, Raman spectra have also been used to determine quantitatively the fraction of the secondary structure in proteins in solution.

VCD, as a new chiroptical technique, has been tested on a number of poly-amino acids including PLL and PLG to assess its feasibility of discriminating different protein chain conformation. It was shown that it indeed has the necessary discrimination and is even better than ECD.<sup>28</sup> It is therefore essential to measure the ROA spectra of these polyamino acids at similar solution environment as that of ECD and conventional Raman to test the ROA sensitivity toward to the secondary structure of these model polypeptides.

We attempted to measure ROA spectra of the PLL and PLG in the aqueous solution environment that produces typical  $\alpha$ -helix,  $\beta$ -sheet and random coil conformations. Good ROA spectra of PLL and PLG in random coil conformation and PLL in  $\alpha$ -helix form were obtained. Unfortunately, under the experimental condition required to keep the  $\beta$ -sheet conformation of PLL (pH=11, T=50°C), the sample becomes glue-like which prevents ROA measurements because of the violent fluctuation of the Raman intensity and the high background generated by the Rayleigh scattering although a good conventional Raman spectrum of the glue-like sample can be obtained. Likewise, we could not record the ROA spectrum of the PLG in  $\alpha$ -helix form because, this time, PLG precipitates at the pH value required to form  $\alpha$ -helix conformation.

Raman and ROA spectra of the PLL and PLG in aqueous solution at pH=3 and pH=10 are presented in Figs. 4.10 and 4.11 respectively. Their Raman and ROA band frequencies and signs, and vibrational assignments, are summarized in Table 4-5.

The ROA spectra of the two model polypeptides in random coil form are rather similar. The most striking ROA features appear in the amide I and amide III regions. Both PLL and PLG show a strong positive ROA peak

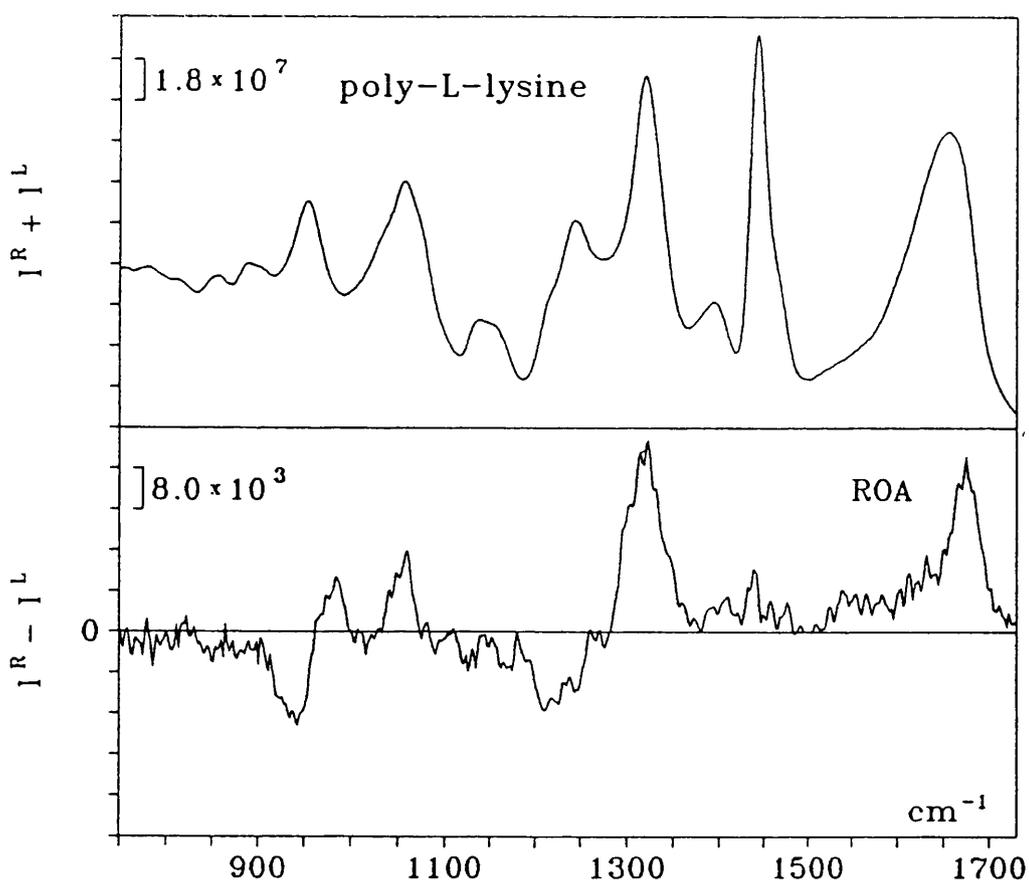


Fig. 4.10 Raman and ROA spectra of PLL in aqueous solution (pH=3).

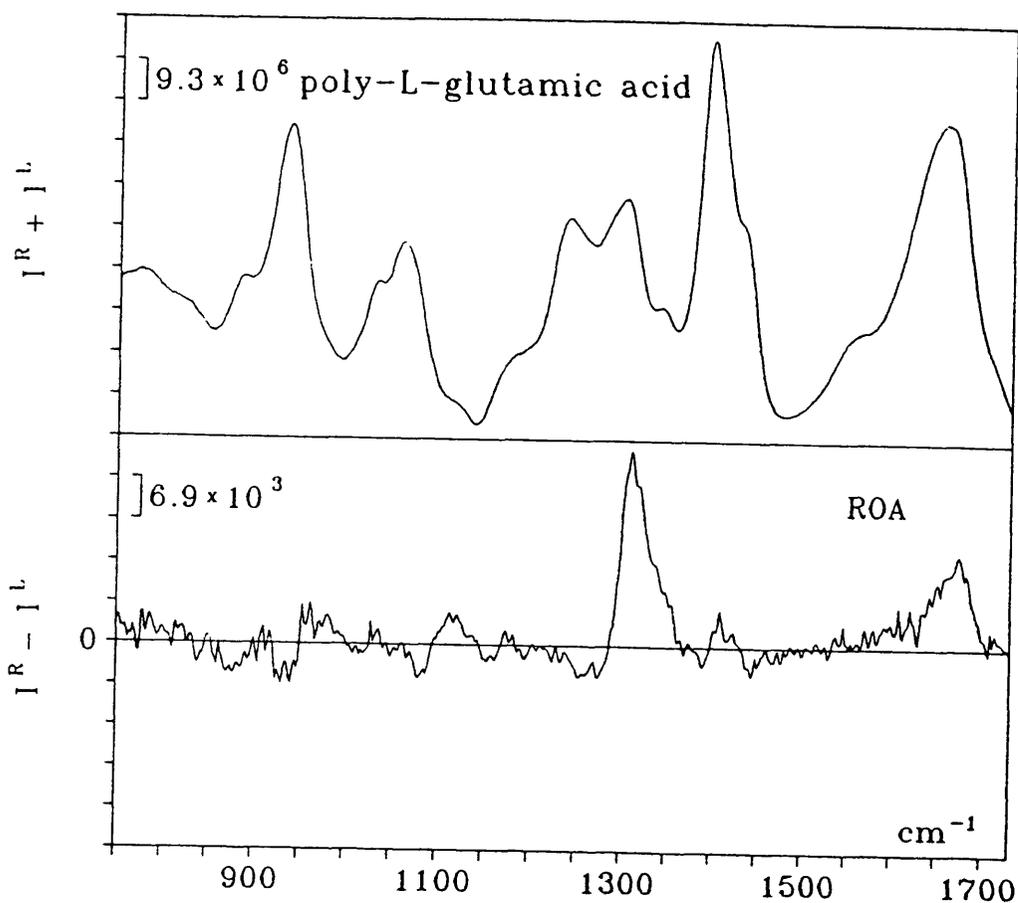


Fig. 4.11 Raman and ROA spectra of PLG in aqueous solution (pH=10).

at  $\sim 1680\text{cm}^{-1}(+)$ , which is apparently related to the amide I C=O stretch vibration. In the amide III region, PLL and PLG give rise to a very strong positive band at  $1320\text{cm}^{-1}(+)$  and  $\sim 1315\text{cm}^{-1}(+)$  and a relatively small negative feature at  $1250\text{cm}^{-1}(-)$  and  $\sim 1260\text{cm}^{-1}(-)$  respectively. These ROA features can be attributed to the extended amide III modes. The positive Raman band at  $\sim 1320\text{cm}^{-1}(+)$  is probably related to the extended delocalized amide III<sup>2</sup> and amide III<sup>3</sup> modes involving the N-H and C-H deformations and the negative feature is due to the amide III<sup>1</sup>, in analogy with the alanyl oligopeptides. PLL shows an additional negative feature at

Table 4-5. ROA band frequencies and signs of poly-L-lysine and poly-L-glutamic acid at different pH and their tentative assignment.

PLL pH=3, T=20°C	PLL pH=11, T=20°C	PGA pH=10, T=20°C	tentative assignment
1680(+)	1660(+)	1680(+)	am I
1450(+)		1470(+)	
1320(+)	1340 (+)	1315 (+)	am III
	1290 (+)		am III
1250(-)	1260 (-)	1260(-)	am III
1210 (-)		1240 (-)	
		1130(+)	
1060(+)	1080 (-)	1080(-)	C <sub>α</sub> -N, C <sub>α</sub> -C stretch
985 (+)		985 (+)	
940 (-)	940 (+)	940(-)	C <sub>α</sub> -C stretch

~ 1210cm<sup>-1</sup>(-). The strong sharp Raman bands at ~ 1445cm<sup>-1</sup> in PLL and ~1410cm<sup>-1</sup> PLG associated with the side group chains do not give significant ROA. The overall resemblance of ROA spectra between the two polypeptides with no significant ROA features from side chain groups implies that the major ROA features of polypeptides arise chiefly from the backbone.

Below 1200cm<sup>-1</sup>, PLL shows a small positive feature at ~ 1060cm<sup>-1</sup>(+) and PLG exhibits a small couplet negative at lower frequency ~ 1080cm<sup>-1</sup>(-) and positive at higher ~ 1130cm<sup>-1</sup>(+). Vibrational modes in this region are believed to be associated with the C-N and C<sub>α</sub>-C stretch. Both of the polypeptides show a negative-positive couplet in the frequency

region between  $900\text{-}1000\text{cm}^{-1}$ . It is centred at  $960\text{cm}^{-1}$  and is negative at lower frequency  $\sim 940\text{cm}^{-1}$ (-) and positive at higher at  $\sim 985\text{cm}^{-1}$ (+). The parent Raman spectra of polypeptides generally show medium strong Raman bands between  $900\text{-}950\text{cm}^{-1}$  when the chain conformation is in random coil or  $\alpha$ -helix. They are considered as chain conformation sensitive bands and are assigned to the backbone  $\text{C}_\alpha\text{-C}$  stretch.<sup>13</sup>

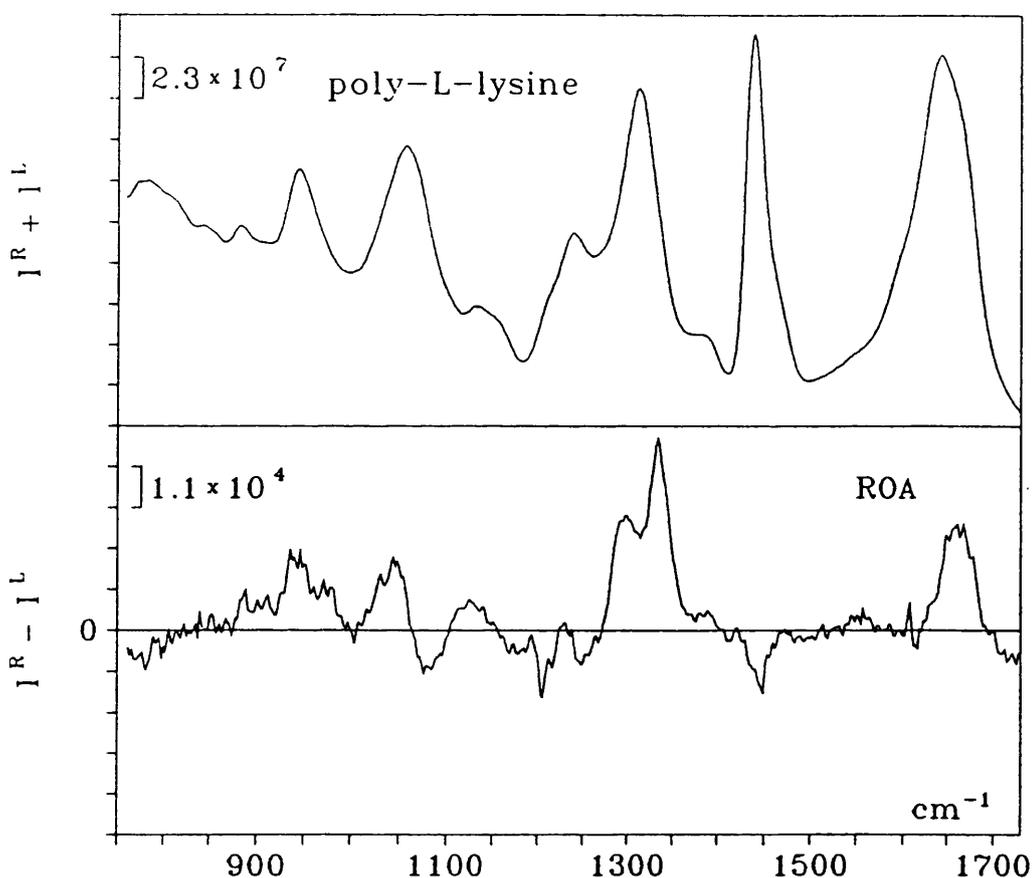


Fig. 4.12 Raman and ROA spectra of PLL in aqueous solution (pH=11).

Fig 4.12 shows the ROA spectrum of poly-L-lysine at pH=11 which differs dramatically from that of at low pH. The positive ROA peak in the amide I region has shifted from  $\sim 1680\text{cm}^{-1}(+)$  at low pH to  $\sim 1660\text{cm}^{-1}(+)$  at high pH but does not change sign. The most important change taken place are in the amide III region. The original single large positive feature at  $1320\text{cm}^{-1}(+)$  is split into two components at  $\sim 1290\text{cm}^{-1}(+)$  and  $\sim 1340\text{cm}^{-1}(+)$ . Moreover, a clear negative-positive couplet appears in the  $\sim 1076-1125\text{cm}^{-1}$  region in  $\alpha$ -helix PLL (pH=11) and the negative feature at  $\sim 940\text{cm}^{-1}(-)$  in random coil PLL (pH=3) has changed sign to positive in  $\alpha$ -helix PLL. These dramatic changes clearly demonstrate that ROA is indeed very sensitive to the polypeptide conformation at different pH. It is intriguing to note that there are remarkable similarity of the ROA band at  $\sim 1340\text{cm}^{-1}(+)$  between PLL and the protein BSA.

#### 4.6. Conclusions

This study has shown that high quality ROA spectra of peptides and polypeptides can now be measured routinely and the information contained in the ROA spectra is rich. Intrinsic absolute configuration of small peptides can be inferred directly from the ROA spectra without any ambiguity. The wide spectral region accessible to ROA measurement with no solvent interference are two advantages over VCD. Thus, ROA and VCD will provide highly complementary information with respect to stereochemical structure. ROA seems to be a very local effect and extremely sensitive to local structural changes, which is similar to VCD although the mechanisms generating ROA and VCD are completely

different.

The preliminary analysis of the Raman and ROA spectra of the di-L-alanine suggests that it could become a good model compound for the analysis of ROA and VCD spectra of peptides and polypeptides and for better understanding the mechanism of ROA generation. Among the many ROA features in the peptides, amide I, amide III, and bands associated with skeleton vibrational modes are probably the most important and useful for protein study. It is worthwhile in the future to carry out a detailed ROA study of the di-L-alanine in the way done by conventional Raman and VCD by precision deuteration of the protons in the molecule to gain a full understanding of the vibrational modes related to the ROA features. Such an effort will no doubt be rewarded in subsequent ROA studies of polypeptides and proteins and will play an important role similar to the model compound N-methylacetamide in the conventional vibrational spectra of peptides and polypeptides.

ROA spectra of the tripeptide for the model  $\beta$ -turn could be used to interpret the similar ROA features in proteins. ROA spectra of the two model polypeptides show that they can be used as models for proteins and their ROA spectra are sensitive to the pH changes. It can be inferred that the dominant ROA features of the two polypeptides arise mainly from the chain backbone with much less contributions from the side chains.

## REFERENCES

1. P. Sutton and J. L. Koenig, *Biopolymers*, **9**, 615, 1970.
2. J. L. Keonig, *J. Poly. Sci. D. Macromolecular Reviews*, **6**, 59, 1972.
3. R. J. Fletterick, Chun-Che Tsai, and R. E. Hughes, *J. Phys. Chem.*, **75**, 918, 1971.
4. M. Goodman, F. Naider and R. Rupp, *Bioorg. Chem.*, **1**, 294, 1971.
5. L. D. Barron, A. R. Gargaro, L. Hecht, P. L. Polavarapu, *Spectrochim. Acta*. **47A**, 1001, 1991.
6. T. Miyazawa, T. Shimanouchi and S. Mizushima, *J. Chem. Phys.*, **29**, 611, 1958.
7. M. R. Oboodi, C. Alva and M. Diem, *J. Phys. Chem.*, **88**, 501, 1984.
8. G. M. Roberts, O. Lee, J. Calienni and M. Diem, *J. Am. Chem. Soc.*, **110**, 1749, 1988.
9. M. Diem, O. Lee and G. M. Roberts, *J. Phys. Chem.*, **96**, 548, 1992.
10. S. Krimm and J. Bandekar, *Adv. Prot. Chem.*, **38**, 181, 1986.
11. L. D. Barron, *Molecular Light Scattering and Optical Activity*, Cambridge University Press, Cambridge, P344, 1982.
12. G. C. Percy and H. S. Stenton, *J. Chem. Soc., Dalton*, 2429, 1976.
13. J. L. Koenig and B. G. Frushour, *Biopolymers*, **11**, 1871, 1972.
14. J. K. Fawcett, N. Camerman and A. Camerman, *Acta Cryst.* **B31**, 658, 1975.
15. W. Qian, J. Bandekar and S. Krimm, *Biopolymers*, **31**, 193, 1991.
16. S. Krimm, J. Bandekar, *Biopolymers*, **19**, 1, 1980.
17. L. L. Reed and P. L. Johnson, *J. Am. Chem. Soc.*, **95**, 7523, 1973.

18. V. M. Naik and S. Krimm, *Internat. J. Pep. Prot. Res.*, **23**, 1, 1984.
19. R. Walter, I. Bernal and L. F. Johnson, *Chemistry and Biology of Peptides*, Ann Arbor Science Publishers, Ann Arbor, Mich., 131, 1972.
20. T. H. Hseu and H. Chang, *Biochimica and Biophysica Acta*, **624**, 340, 1980.
21. J. A. Fox, A. T. Tu, V. J. Hruby and H. I. Mosberg, *Arch. Biochem. Biophys.*, **211**, 618, 1981.
22. P. Lagant, G. Vergoten, G. Fleury and M. H. Loucheux-Lefebvre, *J. Raman Spectrosc.*, **15**, 421, 1984.
23. P. Lagant, G. Vergoten, G. Fleury and M. H. Loucheux-Lefebvre, *Eur. J. Biochem.*, **139**, 137, 1984.
24. P. Lagant, G. Vergoten, G. Fleury and M. H. Loucheux-Lefebvre, *Eur. J. Biochem.*, **139**, 149, 1984.
25. E. R. Blout, F. A. Bovey, M. Goodmans and N. Lotan, *Peptides Polypeptides and Proteins*, John Wiley and Sons, New York, 1974.
26. N. Greenfield and G. D. Fasman, *Biochemistry*, **8**, 4108, 1969.
27. P. R. Carey, *Biochemical Applications of Raman and Resonance Raman Spectroscopies*, Academic Press, New York, p78, 1982.
28. T. A. Keiderling, in 'Practical Fourier Transform Infrared Spectroscopy' edited by J. R. Farraro and K. Krishinan, Academic Press, New York, 203, 1990.

## CHAPTER 5

# Raman optical activity of proteins

### 5.1 Introduction

Proteins are one of the most important classes of biological macromolecules in nature. They play a vital role in life process ranging from transport of oxygen to enzymatic activity. Proteins are composed of polypeptides made by condensation polymerization of amino acids. There are twenty basic amino acids that constitute proteins and all adopt the L-configuration (except glycine which is achiral) and differ from each other by only one side chain. The arrangement of the twenty amino acid residues in a polypeptide chain creates an enormous variety of proteins that have very complicated three dimensional structures and assume numerous functions. To determine the protein structure and to understand the mechanism of protein function are the central themes in protein chemistry.<sup>1</sup>

A protein structure is generally divided into four levels, namely the primary, the secondary, the tertiary and the quaternary structure.<sup>2</sup> The primary structure is the amino acid sequence in a polypeptide chain. The secondary structure means the stable polypeptide chain backbone conformation which is primarily dictated by the intrachain molecular interaction through hydrogen bonding and Van der Waals forces. The

tertiary structure refers to the folding of the polypeptide chain into the final three dimensional shape of the protein and the quaternary is formed by several protein domains adhering together with each such domain containing individual polypeptide chains folded into a certain shape. Except for amino acid sequence, which is determined by chemical degradation method or DNA decoding, all other level structure are measured by spectroscopic techniques such as X-ray diffraction, circular dichroism, vibrational spectroscopy and electron microscopy. The major role played by Raman spectroscopy in protein chemistry is to study the secondary structure of proteins in various states and to probe the molecular interactions.<sup>3</sup> Very recently, multi-dimensional NMR has become an important member in the club of spectroscopic methods to investigate protein structure in solution.<sup>4</sup>

As has been mentioned in chapter 4, the merit of Raman spectroscopy in protein structure determination is that certain Raman band frequencies are related to the peptide linkage of the polypeptide chain and are sensitive to the chain conformation variation. Based on this frequency-conformation relationship and the presumption that the secondary structure in proteins is basically composed of three typical components, namely,  $\alpha$ -helix,  $\beta$ -sheet, and disordered random coil conformation, statistical procedures have been established to estimate semi-quantitatively the compositions of the chain conformation.<sup>3,5</sup> Moreover, characteristic Raman bands of side groups have been used to study the interaction between proteins and other molecules. Resonance Raman spectroscopy can provide subtle information on protein binding site

and protein dynamics when it is operated in a time resolved mode.<sup>6</sup>

However, inspecting the literature of conventional Raman study of proteins reveals that many pitfalls remain. For instance, incomplete assignment of the complicated Raman spectra of many proteins is one of the longstanding problems.<sup>7</sup> Structural information has been only extracted from the band frequency of the Raman spectra; other source of information such as band intensity and shape have not explored as they should be. Moreover, the secondary structure of a protein is far more complicated than the simple three component model. Another important type of secondary structure is reverse turns, which play a significant role in chain folding in globular proteins and in bioactivity.<sup>8</sup> The attempt to determine turn structure by Raman spectra provoked much controversy.<sup>9-11</sup> The concept of random coil conformation is also an unsettled issue because prediction of the fraction of random coil conformation by either Raman and ECD spectra are rather poor.<sup>12</sup>

With the combination of high resolution power from optical activity and the rich vibrational bands, ROA spectra should provide new insights into the protein structure in solution phase since optical activity is a first order function of the molecular geometry. Of course, much more useful information from ROA measurement on proteins are expected. In this chapter, we report the first Raman optical activity study of eight typical proteins and demonstrate the potential utility of ROA in protein structure investigation.

## 5.2 Experimental section

We have chosen eight globular proteins in the initial ROA measurement. They are bovine serum albumin (BSA), insulin (INS), lysozyme from egg white (LZM),  $\alpha$ -lactalbumin ( $\alpha$ -LC) from bovine milk, ribonuclease A (RNS),  $\alpha$ -chymotrypsin ( $\alpha$ -CH),  $\beta$ -lactoglobulin ( $\beta$ -GL), and ovalbumin (OVA). These proteins have been intensively investigated by conventional Raman spectroscopy.<sup>3</sup> They are classified as dominant  $\alpha$ -helix (BSA, INS), dominant  $\beta$ -sheet ( $\alpha$ -CH, RNS), and mixture of  $\alpha$ -helix and  $\beta$ -sheet (LZM,  $\alpha$ -LC,  $\beta$ -GL and OVA) in terms of predominant secondary structure to form the three dimensional shape.<sup>13</sup>

All the eight proteins were purchased with the highest purity available

Table 5.1. The source of proteins studied

Protein	Catalog No.	Source	purity	Company
BSA	05470	bovine serum	1xcry. lyophilized	Fluka
INS	57590	bovine pancrease	—	Fluka
OVA	A-5503	chicken egg	1xcry. lyophilized	Sigma
LYZ	L-6876	chicken egg white	3xcry. lyophilized	Sigma
$\alpha$ -LC	L-5385	bovine milk	lyophilized	Sigma
RNS	83833	bovine pancreas	—	Fluka
$\alpha$ -CH	27270	bovine pancreas	lyophilized	Fluka
$\beta$ -GL	L-0130	bovine milk	3xcry. lyophilized	Sigma

Table 5.2. The secondary structure composition of the eight proteins<sup>a</sup>

	$\alpha$ -helix	$\beta$ -sheet	irregular	methods
BSA	55%	—	45%	ORD in solution
INS	52%	6%	42%	X-ray in crystal
OVA	25%	25%	50%	ORD in solution
LYZ	29~42%	10%	48~62%	X-ray in crystal
$\alpha$ -LC	29~42%	10%	48~62%	ORD in solution
RNS	18%	33~45%	35~45%	X-ray in crystal
$\alpha$ -CH	5%	52%	45%	Raman in solution
$\beta$ -GL	10~20%	30~50%	50%	ORD in solution

<sup>a</sup> adapted from reference 3.

commercially. The sample sources are compiled in Table 5-1. Their secondary structure compositions determined by different spectroscopic methods are summarized in Table 5-2 for the convenience of comparison.

The general sample handling for ROA measurement have been detailed in chapter 3. For proteins, the sample preparation is more demanding. They are generally dissolved in double-distilled water with neutral pH except for insulin and  $\alpha$ -chymotrypsin which were dissolved in pH=2 (KCl+HCL ) and pH=4.7 (acetate) buffer respectively. The protein concentration was prepared at ~ 150mg/ml and the sample solutions filtered with a Millipore membrane filter (0.22 $\mu$ ) to exclude any dust particles. ROA

measurement of proteins are usually carried out on freshly prepared solutions. Some protein solutions showed initially very high fluorescence background. In these cases, the samples were exposed to the laser beam overnight to reduce the background before ROA measurement.

### 5.3 Results and discussion

Figs. 5.1–5.8 show the Raman and ROA spectra of the eight proteins in the spectral region between 1730–750 $\text{cm}^{-1}$ . We were unable to record reliably protein ROA spectra in aqueous solution below 750 $\text{cm}^{-1}$  because of stray light from the high Rayleigh scattering which is even stronger for proteins with high molecular weight and more domains.

To facilitate the assignment and discussion of the ROA features and comparison between similar proteins, we attempt to divide the entire ROA spectra of proteins into three sub-region, namely, the amide I region 1730–1370 $\text{cm}^{-1}$ , the extended amide III region 1370–1200 $\text{cm}^{-1}$  and the skeleton vibrational mode region below 1200 $\text{cm}^{-1}$ .

**1730–1370 $\text{cm}^{-1}$  region.** This region covers the amide I and amide II bands plus some ROA features arising from aromatic amino acid residues such as tryptophan and methyl groups. The dominant ROA feature in most proteins is a negative-positive couplet from the amide I C=O stretch coordinate. The amide II band does not exhibit any significant ROA features; however, it gives rise to a positive ROA peak at the lower frequency amide II' region when proteins are dissolved in D<sub>2</sub>O for N-deuteration which replace the proton in N-H group by a deuterium.

Tryptophan and tyrosine residues can show striking ROA features over the entire spectrum; particularly in this region, strong tryptophan bands occur at  $\sim 1620\text{cm}^{-1}$ ,  $\sim 1580\text{cm}^{-1}$ ,  $1550\text{cm}^{-1}$  in the ROA spectra of proteins containing larger numbers of tryptophan residues.

**$1370\text{--}1200\text{cm}^{-1}$  (the extended amide III region).** Our division of the amide III region differ from the conventional amide III area which cover only a rather narrow spectral range between  $1230\text{--}1290\text{cm}^{-1}$ .<sup>3</sup> We propose that the amide III region should extend up to  $1370\text{cm}^{-1}$  and call it the 'extended amide III' region for a number of arguments. First of all, we have observed that ROA spectra of peptides, polypeptides and proteins give rise to ample ROA features in the spectral range above  $\sim 1300\text{cm}^{-1}$  which are all associated with the amide III vibrational modes. Secondly, recent detailed Raman and VCD studies of the dipeptide di-L-alanine disclosed that the so-called amide III band is actually composed of three basic components, one of which (amide III<sup>3</sup>) exists beyond the conventional amide III region. Thirdly, normal mode calculation of polypeptide chains suggested that vibrational Raman bands presented above  $1300\text{cm}^{-1}$  are often associated with reverse turn structure of peptides and proteins.<sup>9</sup> As a matter of fact, deuteration studies of many proteins show that many bands between  $1300\text{--}1370\text{cm}^{-1}$  are more or less related to the N-H group because they either disappear or change frequency upon N-deuteration.<sup>5</sup> Finally, vibrational Raman bands of proteins between  $1300\text{--}1370\text{cm}^{-1}$  were either poorly assigned or simply ignored in previous Raman studies possibly owing to the poor understanding of the Raman bands in this region.<sup>3</sup>

**Below  $1200\text{cm}^{-1}$  ( the skeleton region ).** Below  $1200\text{cm}^{-1}$ , ROA features are relatively weak in contrast to the amide I and amide III regions. Most of the vibrational bands are delocalized. Nevertheless, ROA bands associated with the  $\text{C}_\alpha\text{-N}$  stretch and the  $\text{C}_\alpha\text{-C}$  stretch are discernible. The most significant ROA bands are a negative and positive ROA couplet at  $\sim 1130\text{cm}^{-1}$ . Another interesting ROA feature is a positive peak at  $930\text{-}960\text{cm}^{-1}$  in proteins having high fraction of  $\alpha$ -helix component. ROA features from tryptophan and tyrosine can be seen in the lower frequency region. Thus, tryptophan shows a small negative feature at  $\sim 760\text{cm}^{-1}$  and tyrosine exposes small ROA bands at  $820\text{-}870\text{cm}^{-1}$ .

### **5.3.1 Bovine serum albumin, insulin and ovalbumin**

ROA spectra of these three proteins are grouped together for comparison and discussion since they contain a relatively high fraction of  $\alpha$ -helix component in their secondary structure. BSA has over 55% amino acid residues in the  $\alpha$ -helix segments and the rest of residues are in disordered structure. A striking feature of the BSA structure is that it possesses a large amount of cystine residues (34) which compose 17 disulfide bridge to crosslink the single polypeptide chain and form many double loops (Fig. 5.1).<sup>14</sup> INS is well-known for its many firsts in protein chemistry: for example, it was the first sequenced and synthesized protein, and the first protein used as therapeutic drug.<sup>15</sup> It is an excellent model protein for protein structure study since the absence of tryptophan residues reduces the complexity of the spectrum. Its three dimensional structure is well-established and 50% of its residues occur in  $\alpha$ -helix.

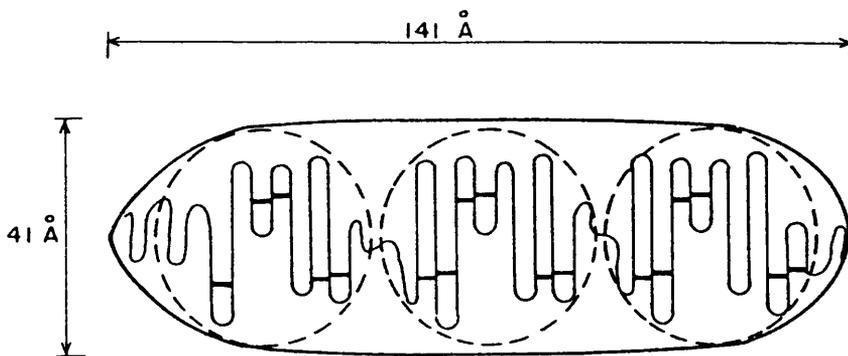


Fig. 5.1 Model of the BSA molecule. The S-S bridges are indicated by the dark bar.

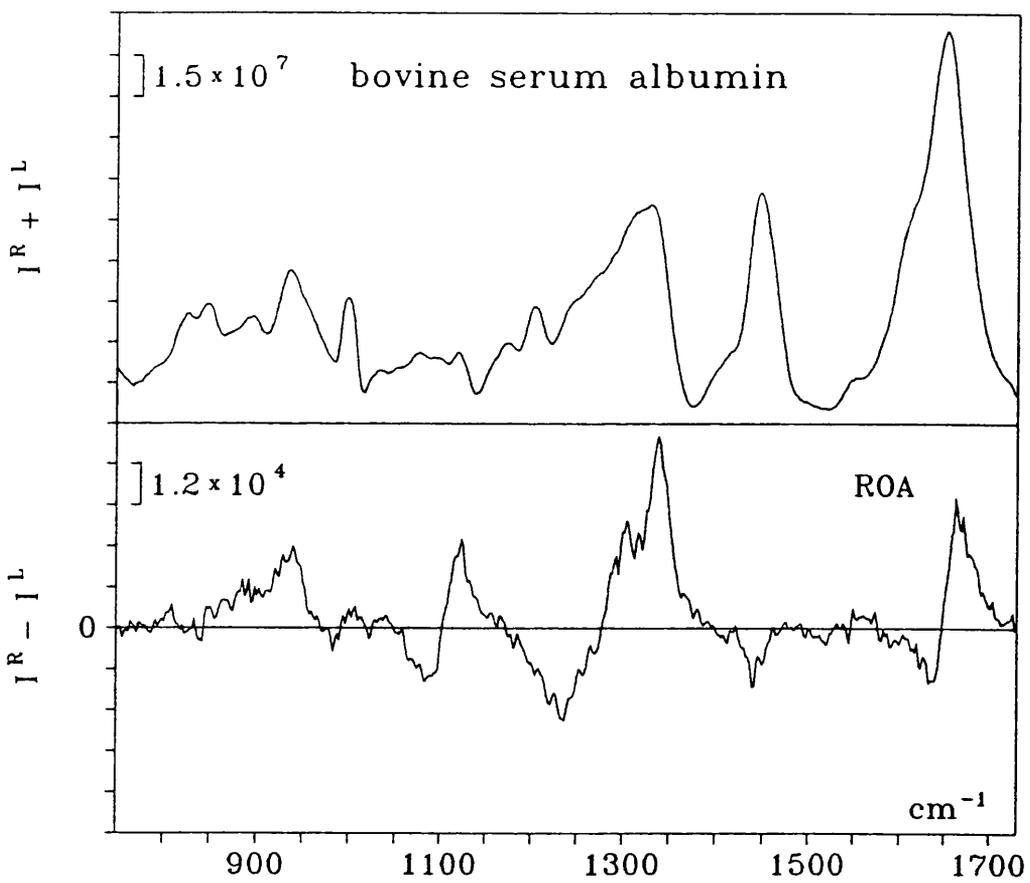


Fig. 5.2 Raman and ROA spectra of BSA in water.

ovalbumin is a globular protein in egg and is important protein in the food industry. Its secondary structure is measured by ORD and comprises 25%  $\alpha$ -helix and 25%  $\beta$ -sheet, the rest being random coil. Raman spectra give rise to similar results of the secondary composition.<sup>3</sup>

ROA spectra of the three proteins are relatively simple compared to other proteins since BSA has merely two tryptophan residues and none in INS. In the amide I region, there is a negative-positive ROA couplet in BSA and OVA but not in INS (Figs. 5.2-5.4). It is probable that the negative part of the couplet in INS is superimposed by some positive ROA features from tyrosine and phenylalanine residues, thus cancelling each other. At  $\sim 1440\text{cm}^{-1}$ , all the three proteins have a small negative feature which could be due to the side chain methylene and methyl groups.

The major ROA features, similar to all other proteins, are concentrated in the extended amide III region. The most salient ROA feature is a strong positive peak at  $\sim 1340\text{cm}^{-1}(+)$  together with a positive shoulder at  $1300\text{cm}^{-1}(+)$  in BSA. Insulin only shows a broad positive shoulder at  $1340\text{cm}^{-1}(+)$  and a positive peak at  $\sim 1310\text{cm}^{-1}(+)$ . OVA exhibits a strong positive peak at  $\sim 1310\text{cm}^{-1}(+)$  and two positive shoulders at  $\sim 1340\text{cm}^{-1}(+)$  and  $1295\text{cm}^{-1}(+)$ . All the three proteins show a broad negative band centred at  $\sim 1230\text{cm}^{-1}(-)$ ,  $1240\text{cm}^{-1}(-)$  and  $\sim 1223\text{cm}^{-1}(-)$  respectively.

Raman bands in this region are related to the complicated amide III modes which are believed to depend on both polypeptide chain conformation and the side groups.<sup>3</sup> The origin of the positive ROA band at  $\sim 1340\text{cm}^{-1}(+)$  in BSA can be attributed to the N-H groups since the band

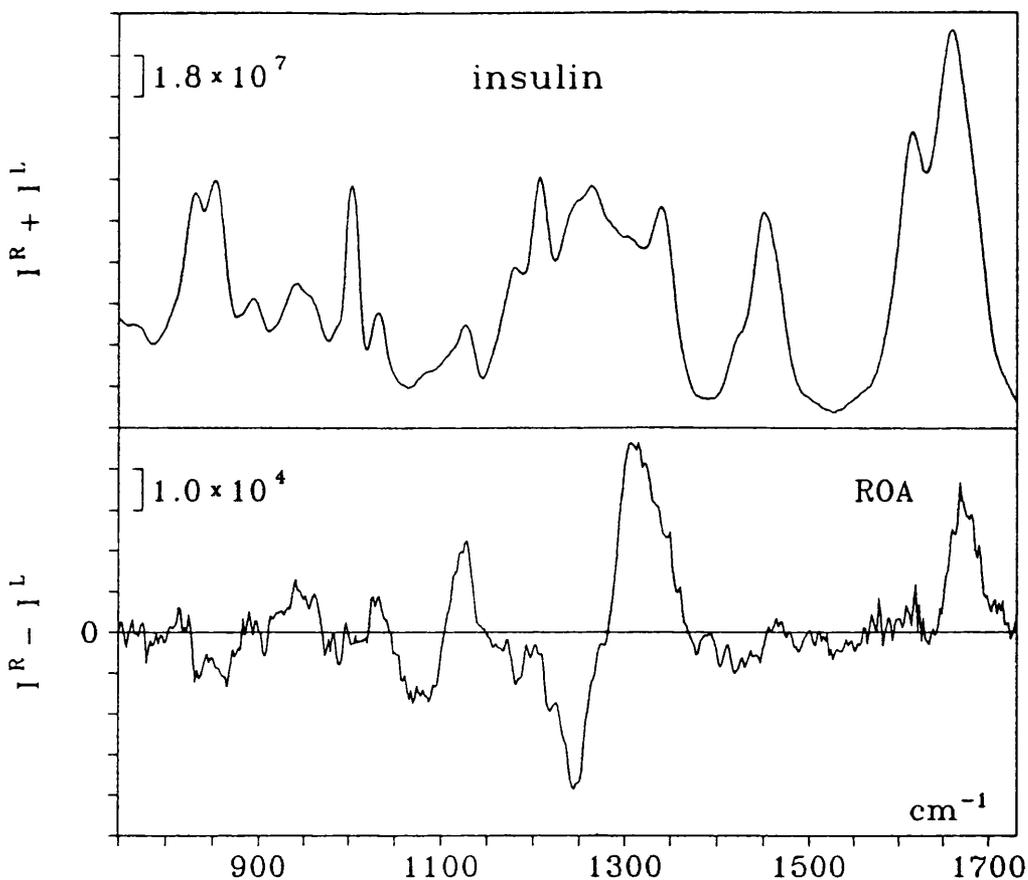


Fig. 5.2 Raman and ROA spectra of insulin at pH=2 aqueous solution.

disappears upon N-deuteration in  $D_2O$ .<sup>16</sup> On the contrary, the next positive shoulder band at  $\sim 1300\text{cm}^{-1}(+)$  remains unchanged in  $D_2O$ . ROA features of dipeptides in this region have been assigned to the amide III<sup>3</sup> and amide III<sup>2</sup> mode for  $\sim 1340\text{cm}^{-1}(+)$  and  $\sim 1300\text{cm}^{-1}(+)$  bands respectively. The ROA peaks at  $\sim 1340\text{cm}^{-1}(+)$  and  $\sim 1300\text{cm}^{-1}(+)$  in these proteins could probably be attributed to the similar amide III<sup>3</sup> and amide III<sup>2</sup> modes since ROA is a very local effect.

The remarkable difference of the amide III ROA features between the two proteins BSA and INS could be due to the difference of their detailed

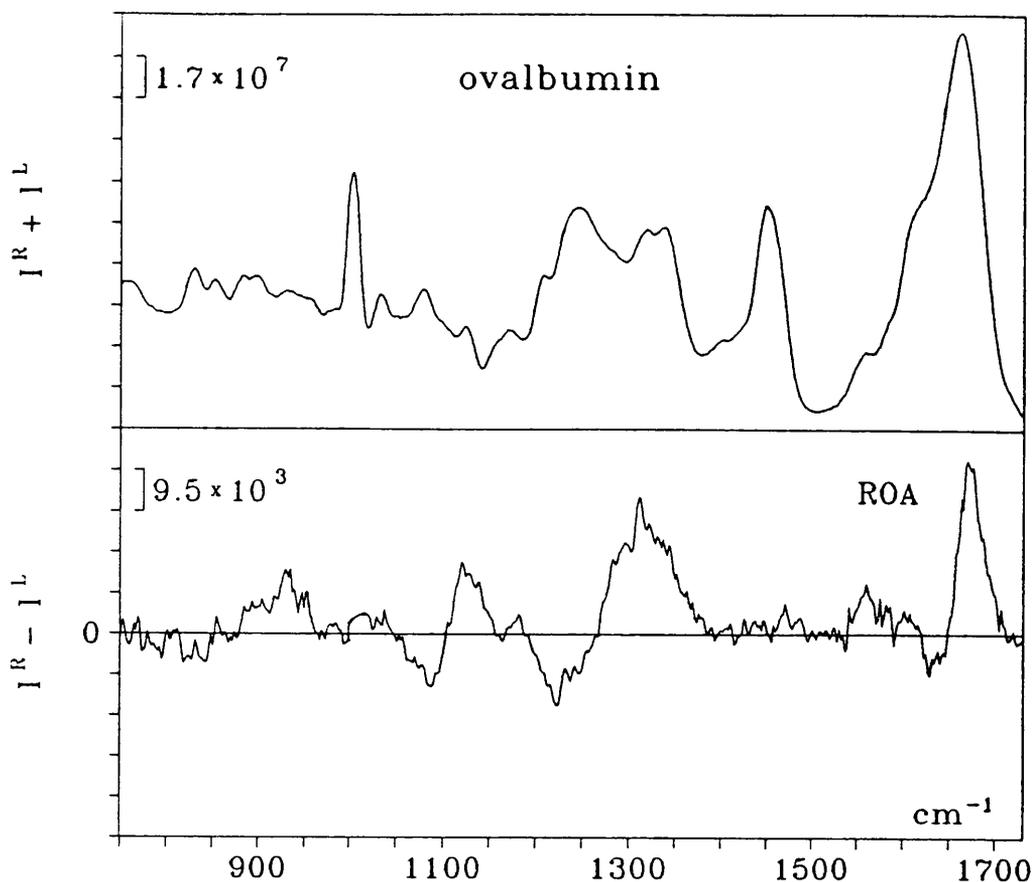


Fig. 5.4 Raman and ROA spectra of ovalbumin in water.

secondary structure. Although both BSA and INS contain similar high fraction of  $\alpha$ -helix components, the total amino acid residues of the two proteins is ten times different. As it has now been confirmed from high resolution X-ray results,<sup>17</sup> serum albumin indeed has a large number of double loops and many tight reverse turns whereas INS has merely four reverse  $\beta$ -turns.<sup>9</sup> There is an abnormally high content of proline residues (27) in BSA, and half of them occurring at the turn positions.<sup>14</sup> It has long been recognized that L-proline and L-cystine are the two most important amino acid residues to promote turn structure in globular

proteins.<sup>13</sup> Indeed, a huge positive ROA peak at  $\sim 1338\text{cm}^{-1}$  (+) occur in the ROA spectrum of L-proline. We therefore suggest that the prominent positive ROA peak at  $\sim 1340\text{cm}^{-1}$ (+) in BSA is associated with the reverse turns and loops of the polypeptide chain.

Below  $1200\text{cm}^{-1}$ , ROA spectra of all three proteins are rather simple. They all show a negative-positive ROA couplet centered at  $\sim 1100\text{cm}^{-1}$ , negative at  $\sim 1080\text{cm}^{-1}$  and positive at  $\sim 1130\text{cm}^{-1}$  but with different

Table 5-3 Characteristic ROA band frequency of  $\alpha$ -helix of polypeptide chain.

	Raman ( $\text{cm}^{-1}$ )	ROA ( $\text{cm}^{-1}$ )	$\alpha$ -helix content <sup>a</sup>
BSA	938	936 (+)	55%
INS	942	943 (+)	52%
OVA	930	930 (+)	25%
$\beta$ -GL	935	939 (+)	10~20%
$\alpha$ -LC	930	930 (+)	29~42%
LYZ	930	930 (+)	29~42%
PLL	945	940 (+)	>90%

a. data adapted from reference 3.

magnitudes. The couplet in BSA is the strongest and INS the second.

Another intriguing ROA feature in this group of proteins is that they all show a positive ROA peak at  $\sim 930\text{cm}^{-1}$ (+) which is absent in all

proteins that are either  $\beta$ -sheet dominant or with a marginal magnitude in proteins with low  $\alpha$ -helix fraction (Table 5-3). This indicates that the positive ROA band could be contributed by  $\alpha$ -helix components. Conventional Raman spectra of polypeptides and proteins with high content of  $\alpha$ -helix showed a band at  $\sim 930\text{cm}^{-1}$  which had been revealed to be related to the  $\text{C}_\alpha\text{-C}$  and  $\text{C-N}$  backbone stretching mode by normal mode calculation of the polypeptide chain.<sup>5</sup>

Other ROA features discernible in the lower frequency region include small ROA bands associated with tyrosine residues at  $870\text{-}830\text{cm}^{-1}$  and a negative ROA peak at  $\sim 760\text{cm}^{-1}$ (-) which is apparently related to tryptophan. It is worthwhile to point out that there is a small negative ROA feature at  $\sim 670\text{cm}^{-1}$ (-) in BSA which definitely belong to the C-S stretch mode of the disulfide bridge since it is an isolated band in Raman spectra and BSA has a large amount of S-S bridges (not shown here).

### 5.3.2 Lysozyme and $\alpha$ -lactalbumin

Lysozyme and  $\alpha$ -lactalbumin are an ideal pair of proteins for comparison. Lysozyme is the first enzyme to have its three dimensional structure determined by X-ray diffraction<sup>18</sup> and became a model protein in studies of the relationship between structure and enzyme activity.<sup>19</sup>  $\alpha$ -lactalbumin is a milk protein that has become an excellent model protein for the study of calcium ion binding.<sup>20</sup> The two proteins are structurally analogous. Not only their amino acid sequences are in common for 40%, but also their secondary and tertiary structure are known to be

very similar from X-ray diffraction. Both of them have approximately 30%  $\alpha$ -helix and 40%  $\beta$ -sheet. There are voluminous parallel comparison studies on these two proteins.<sup>21</sup> Conventional Raman has also been conducted to investigate the similarity and discrepancy in chain conformation and changes upon denaturation and pH changes.<sup>22</sup>

It can be seen from Figs. 5.5 and 5.6 that the ROA spectra of the two proteins are very similar in the amide I region. Both show a negative-positive ROA couplet associated with amide I band but the magnitude of the couplet is slightly different, being larger in lysozyme than in  $\alpha$ -lactalbumin. A positive ROA peak at  $\sim 1620\text{cm}^{-1}$  occurs in both proteins which is attributed to the tryptophan residues (LZM and  $\alpha$ -LC possess 6 and 5 tryptophan residues respectively). There is a small negative-positive couplet at  $\sim 1460\text{cm}^{-1}$  which is apparently due to the methyl and methylene groups in side chains.

In contrast to the great similarity in the amide I region, ROA spectra of the two proteins exhibit remarkable differences in the extended amide III region. LZM exhibits three well-resolved positive peaks at  $\sim 1360\text{cm}^{-1}(+)$ ,  $\sim 1340\text{cm}^{-1}(+)$ , and  $\sim 1300\text{cm}^{-1}(+)$  plus a broad negative feature at  $\sim 1250\text{cm}^{-1}(-)$ ; whereas  $\alpha$ -LC displays two very sharp positive peaks at  $\sim 1340\text{cm}^{-1}(+)$  and  $\sim 1310\text{cm}^{-1}(+)$  and a broad negative feature at  $\sim 1250\text{cm}^{-1}(-)$ . The most striking difference is the ROA intensity of the positive band at  $\sim 1340\text{cm}^{-1}(+)$  which is almost twice as strong in  $\alpha$ -LC as that in LZM.

The ROA bands at  $\sim 1300\text{cm}^{-1}(+)$ ,  $\sim 1310\text{cm}^{-1}(+)$  in LZM and  $\alpha$ -LC can

be attributed to the amide III<sup>2</sup> mode since they remain unchanged after N-deuteration. The most interesting ROA signal in this region is the sharp positive peak at  $\sim 1340\text{cm}^{-1}$ (+). The ROA feature can be assigned to the amide III<sup>3</sup> mode arising from the N-H/C-H coupling in the polypeptide chain. N-deuteration reveals that the ROA intensity of this positive ROA band is reduced dramatically but retains an identifiable positive peak. This

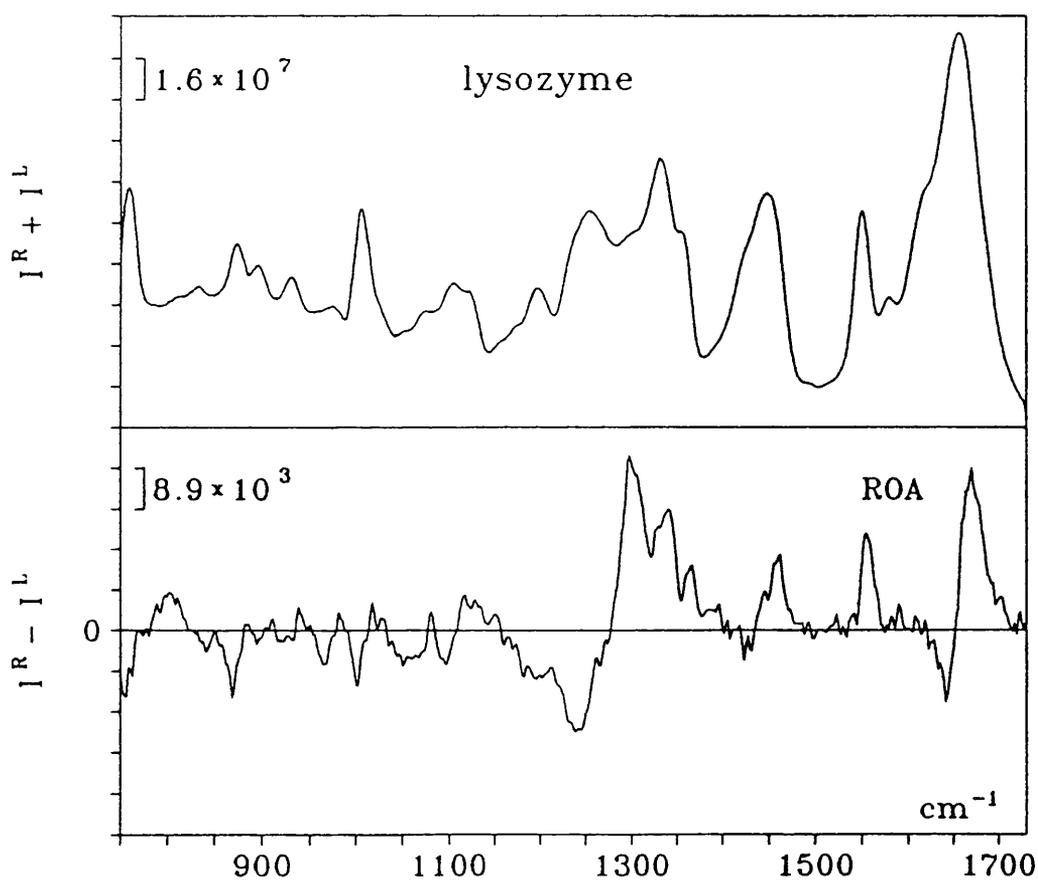


Fig. 5.5 Raman and ROA spectra of lysozyme in water.

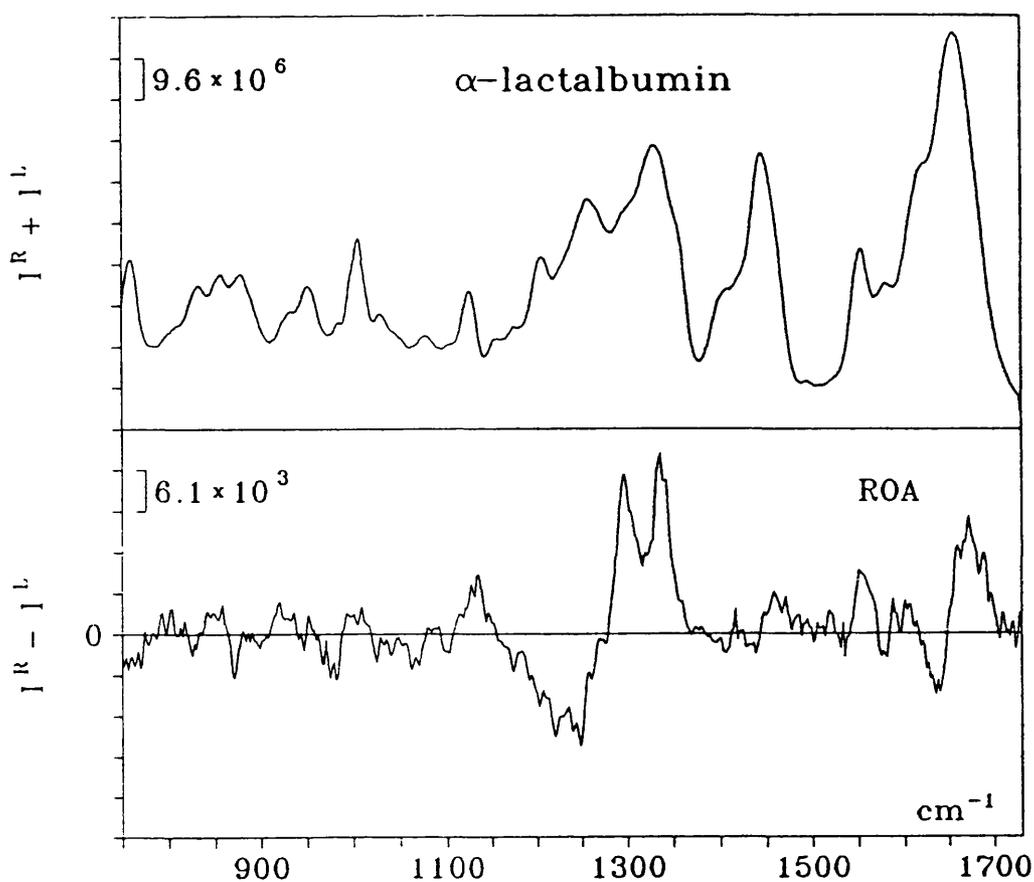


Fig. 5.6 Raman and ROA spectra of  $\alpha$ -lactalbumin in water.

confirms that the assignment of the band to amide III<sup>3</sup> is appropriate but also suggests that there are other components that are responsible for this feature. The most likely contribution is from tryptophan residues since tryptophan shows a strong Raman doublet at this frequency.<sup>23</sup> This interpretation is further supported by the ROA spectra of BSA, which has also a very strong positive band at  $\sim 1340\text{cm}^{-1}(+)$  but it disappears completely upon N-deuteration (remember that BSA contains merely two tryptophan residues in its total 582 amino acid residues).

It is intriguing to note the significant ROA intensity difference of the  $\sim 1340\text{cm}^{-1}$  (+) band between the two proteins. This apparently can not be attributed to the different number of the tryptophan residues since LZM has more tryptophans than  $\alpha$ -LC, nor to the composition of their secondary structure since they are virtually the same. It is therefore considered that this ROA intensity difference is due to the difference of the rigidity or flexibility of the chain conformation. Again loop or turn structure is implied since they are usually more rigid than either  $\alpha$ -helix or  $\beta$ -strand. As it will be discussed late in the chapter on carbohydrate ROA, we have found that the ROA  $\Delta$ -value is extremely sensitive to the rigidity or flexibility of the molecular conformation of the monosaccharides. The more rigid the molecule, the stronger the ROA intensity for similar monosaccharides. The ROA  $\Delta$ -value could be one or two orders of magnitude different between monosaccharide stereoisomers. It has long been recognized that the ROA intensity of crystals could be ten times larger than pure liquids.<sup>24</sup> As is well known,  $\alpha$ -LC is a calcium binding protein and has a rather more compact three dimensional structure than LZM.<sup>21</sup> The intensity difference of the  $\sim 1340\text{cm}^{-1}$ (+) band between lysozyme and  $\alpha$ -lactalbumin may thus be attributed to their different overall structural compactness. It is probably associated with the calcium binding site which is composed of a specific tight turn structure involving both side chain group and main chain backbone,<sup>21</sup> whereas lysozyme does not have such a specific site for calcium binding.

This hypothesis gained more support from the ROA spectra of  $\alpha$ -LC at low pH and of a calcium free sample (Sigma Type III  $\alpha$ -lactalbumin). Upon

lowering the pH of the solution to pH=2, the ROA spectrum of  $\alpha$ -LC changes drastically. The two well-resolved ROA peaks now reduced to one broad positive band with much weaker intensity. The ROA spectrum of the calcium free  $\alpha$ -LC shows overall similar ROA features to that of calcium-bound but the two ROA peaks are now of reduced intensity and less resolved. We will discuss it later in more detail in section 5.3.4.

The ROA spectra of the two proteins in the region below  $1200\text{cm}^{-1}$  are rather similar. A clear positive band at  $\sim 1130\text{cm}^{-1}(+)$  can be seen in both proteins which is related to the  $C_{\alpha}$ -N stretch coordinate. This ROA feature remains in  $D_2O$ . Other ROA features in this region include a small negative band at  $\sim 870\text{cm}^{-1}(-)$  and a quite strong negative band at  $\sim 760\text{cm}^{-1}(-)$ . The former is apparently due to the tyrosine residues and the latter is contributed by the tryptophans. There are some other weak ROA features over the skeleton vibrational region but none of them are significant at the present signal to noise level. Of course, with the improvement of spectral quality, these weak ROA features could be informative about the chain backbone.

### 5.3.3 Ribonuclease A, $\alpha$ -chymotrypsin and $\beta$ -lactoglobulin

Ribonuclease A,  $\alpha$ -chymotrypsin and  $\beta$ -lactoglobulin are classified as high  $\beta$ -sheet proteins since they contain substantial  $\beta$ -sheet component and only a very small fraction of  $\alpha$ -helix. Among the three proteins, ribonuclease A has merely a few residues at the C-terminal in  $\alpha$ -helix and no perfect  $\alpha$ -helical segments. Most of the polypeptide chain is strongly

extended and the adjacent  $\beta$ -strands are connected with reverse turns.<sup>17</sup> X-ray diffraction disclosed a high amount of the extended  $\beta$ -structure in  $\alpha$ -chymotrypsin.<sup>19</sup> With the exception of eight residues at the C-terminal, which form a short section of  $\alpha$ -helix, the chains tend to be fully extended. The secondary structure of  $\beta$ -lactoglobulin was suggested to be similar to that of ribonuclease.<sup>25</sup> It contains about 30-40%  $\beta$ -sheet, 10-20%  $\alpha$ -helix and the rest is disordered fraction. The polypeptide chain is arranged in a total of nine  $\beta$ -strands and merely one  $\alpha$ -helix segment.

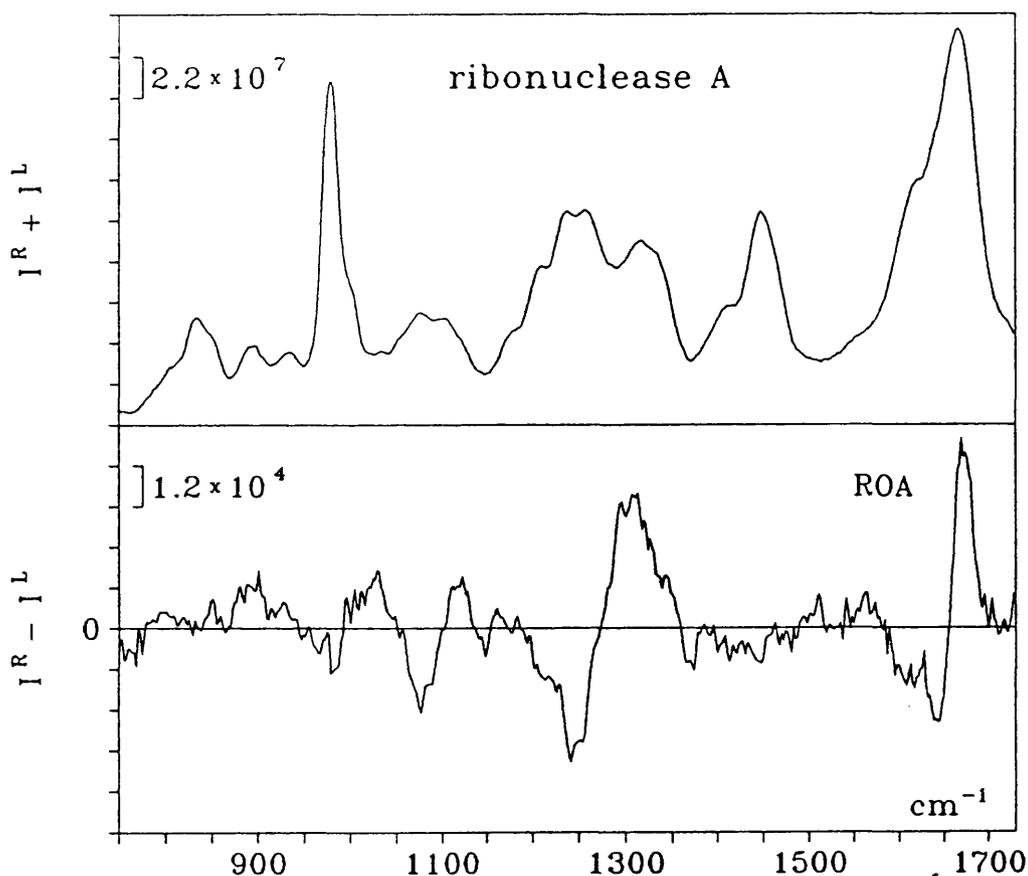


Fig. 5.7 Raman and ROA spectra of ribonuclease A in water.

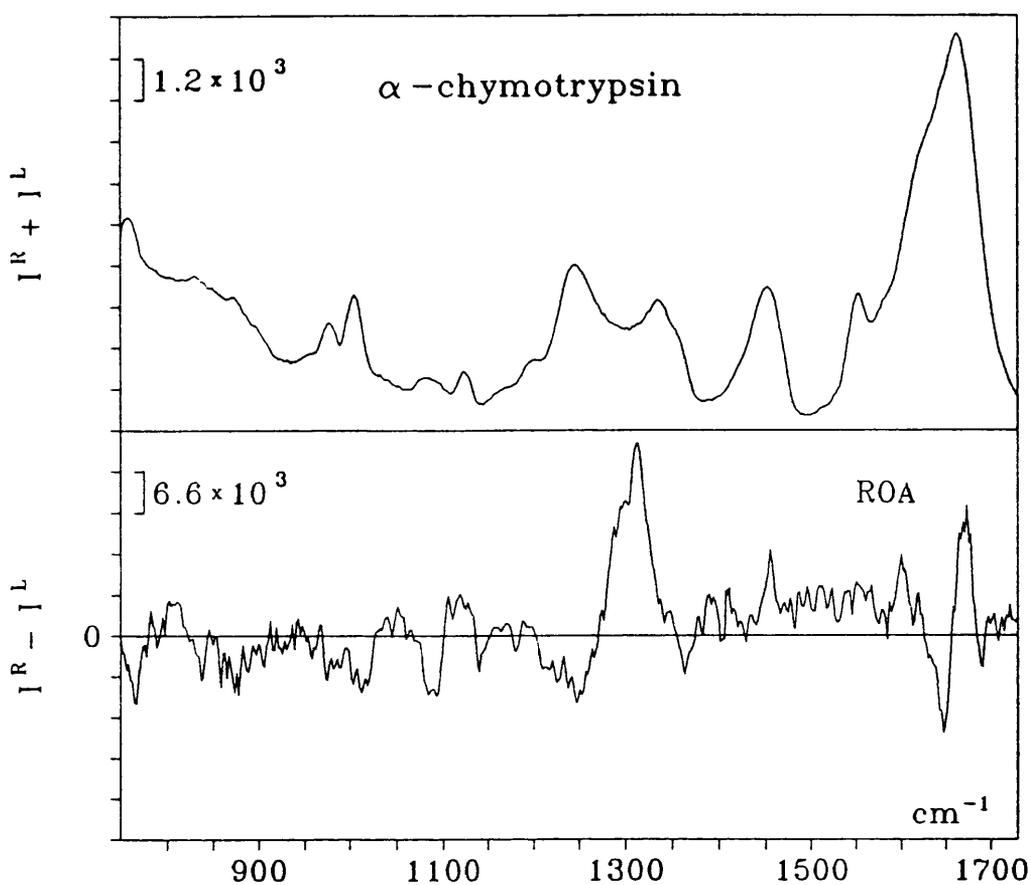


Fig. 5.8 Raman and ROA spectra of  $\alpha$ -chymotrypsin at pH=4.7 aqueous solution.

All the three proteins show a negative-positive couplet in the amide I region in their ROA spectra which is related to the C=O stretch (Figs.5.7-5.9).  $\beta$ -lactoglobulin shows a small positive feature at  $\sim 1580\text{cm}^{-1}$  which is probably due to the tryptophan or tyrosine residues since there are no corresponding features in ribonuclease A which does not have any tryptophan.

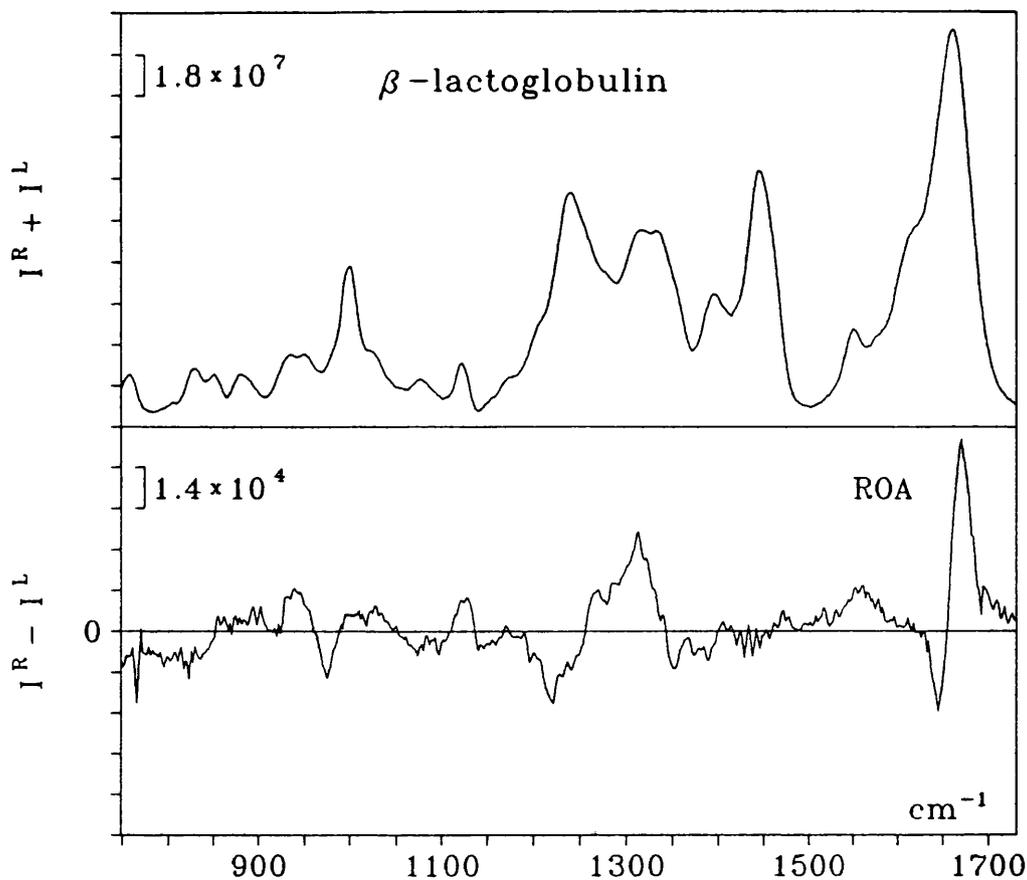


Fig. 5.9 Raman and ROA spectra of  $\beta$ -lactoglobulin in water.

The most significant ROA features common in all the three proteins is that they all show a relatively strong positive peak at  $\sim 1310\text{cm}^{-1}(+)$ . Thus, ribonuclease A gives a strong positive peak at  $\sim 1310\text{cm}^{-1}(+)$  and a positive shoulder at  $\sim 1340\text{cm}^{-1}(+)$ . It is surprising that  $\alpha$ -chymotrypsin exhibits an extraordinary strong positive peak at  $\sim 1310\text{cm}^{-1}(+)$  accompanied by a positive shoulder to lower frequency at  $\sim 1290\text{cm}^{-1}(+)$ .  $\beta$ -lactoglobulin exposes a positive peak at  $\sim 1313\text{cm}^{-1}(+)$  and a positive shoulder to lower frequency at  $\sim 1270\text{cm}^{-1}(+)$ . All the three proteins also show a negative feature between  $1270\text{--}1200\text{cm}^{-1}$ . However, the central negative peak

frequencies are  $\sim 1240\text{cm}^{-1}$  (RNS),  $\sim 1245\text{cm}^{-1}$  ( $\alpha$ -CH) and  $\sim 1220\text{cm}^{-1}$  ( $\beta$ -GLB) respectively. Another intriguing common ROA feature in this group of proteins is a negative peak at  $\sim 1360\text{cm}^{-1}$ (-). None of the other proteins measured so far show such an ROA peak at this frequency.

The vibrational origins of the ROA feature at  $\sim 1310\text{cm}^{-1}$ , should be similar to the other proteins, namely associated with the amide III modes. The magnitude of the positive peak at  $\sim 1310\text{cm}^{-1}$ (+) is extraordinarily large in  $\alpha$ -chymotrypsin, which contains a large fraction of  $\beta$ -sheet components (52%). This indicates that the band could be related to the  $\beta$ -sheet structure. One of the most characteristic chain arrangement of proteins in this group is that, apart from forming substantial  $\beta$ -strands, all these  $\beta$ -strands are connected by rather tight reverse turns with the neighbouring strands being formed from adjacent segments of the polypeptide chain.<sup>13</sup>

Likewise, the negative ROA feature at  $\sim 1360\text{cm}^{-1}$ (-) can also be proposed to be associated with the  $\beta$ -turn structure. It is probably related to  $\beta$ -turns at the crossover and adjacent connections of the  $\beta$ -strand. It is interesting to recall that the model  $\beta$ -turn tripeptide indeed shows a small negative feature at this frequency, although at present we are not able to assign this vibrational band to a definite vibrational mode. Nevertheless, normal mode calculation do predict that the band is contributed by the N-H bending from proline residue and the  $\text{CH}_2$  deformation of the leucine residue.<sup>26</sup> As is well known, proline residues play a vital role in the folding of globular proteins and occurs often at the reverse turn position.<sup>8</sup>

### 5.3.4 $\alpha$ -lactalbumin at low pH and upon calcium depletion

$\alpha$ -lactalbumin offers an excellent protein to test the sensitivity of ROA to probe the protein structural change. Conventional Raman spectroscopy has been used to probe the conformational variation during pH change and to detect the conformational difference in crystal and solution phase.<sup>22</sup> FTIR spectroscopy has been used to investigate the secondary structure of  $\alpha$ -lactalbumin<sup>24</sup> and VCD of the amide I region has also been employed to discriminate the structural difference of  $\alpha$ -LC between solid and solution phase.<sup>25</sup> It would be interesting to see whether Raman optical activity is sensitive to the structural changes during the pH variation and calcium depletion.

The ROA spectrum of the  $\alpha$ -LC at low pH was obtained after adjustment of the pH of the aqueous solution of the native  $\alpha$ -LC. The spectral region measured is restricted to above  $1000\text{cm}^{-1}$  because the solution at low pH show very bright Rayleigh scattering, possibly due to the high charge content in the protein chains upon lowering pH (we have observed that polyamino acids and some proteins, for instance, insulin exhibit much stronger Rayleigh scattering at low pH aqueous solution).

Figs. 5.10-5.11 show ROA spectra of the two samples at spectral range between  $1750\text{-}750\text{cm}^{-1}$  for the calcium depleted  $\alpha$ -LC (Sigma Type III  $\alpha$ -lactalbumin) and  $1750\text{-}1000\text{cm}^{-1}$  for the native protein at pH=2 respectively.

The most striking changes of the ROA spectra upon removal of the calcium is in the amide I and III region. In contrast to the two

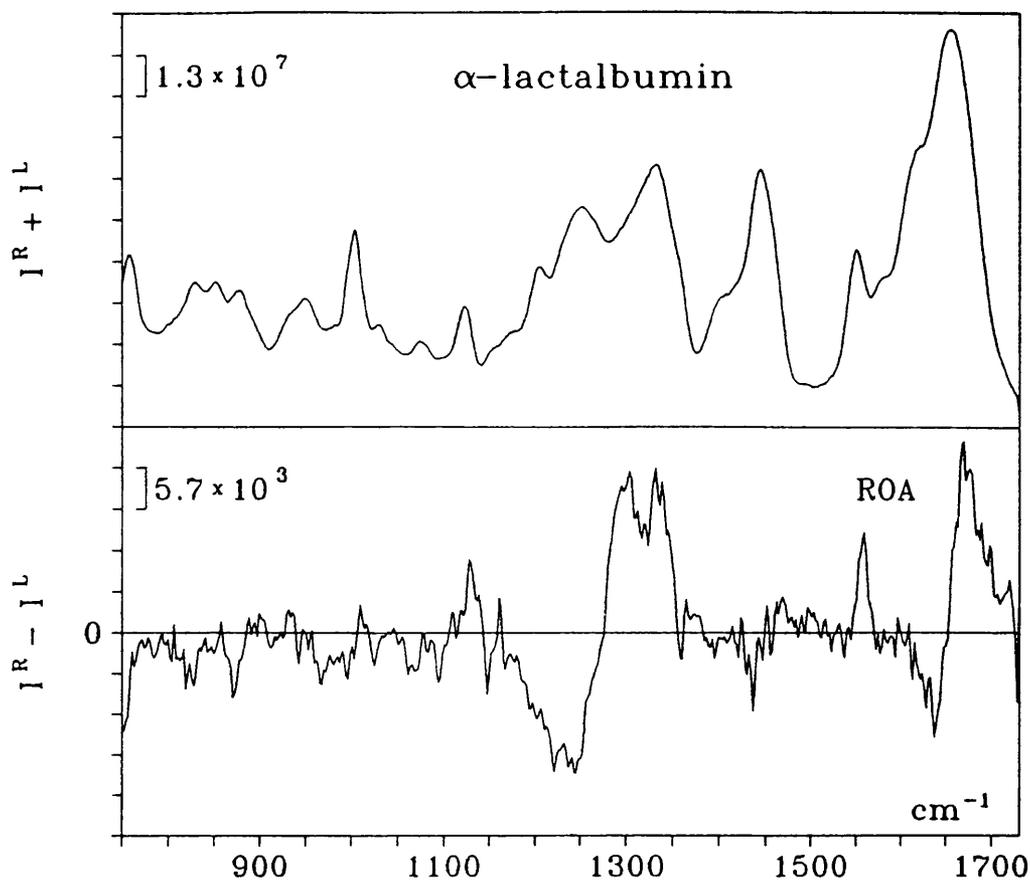


Fig. 5.10 Raman and ROA spectra of  $\alpha$ -lactalbumin (calcium free) in water.

well-resolved sharp positive peaks at  $\sim 1340\text{cm}^{-1}(+)$ ,  $\sim 1310\text{cm}^{-1}(+)$  in the native  $\alpha$ -LC, the two bands show reduced intensity and resolution and coalesce together on the removal of the calcium ions. A rough estimate of the ROA  $\Delta$  value of the  $\sim 1340\text{cm}^{-1}(+)$  band reveals that the ROA intensity dropped nearly to half of the original value when the calcium is depleted. Moreover, the relative magnitude of the ROA intensity between the amide I and amide III have also altered and the overall SNR of the ROA spectra is smaller in the calcium free sample when it is measured with the same acquisition time.

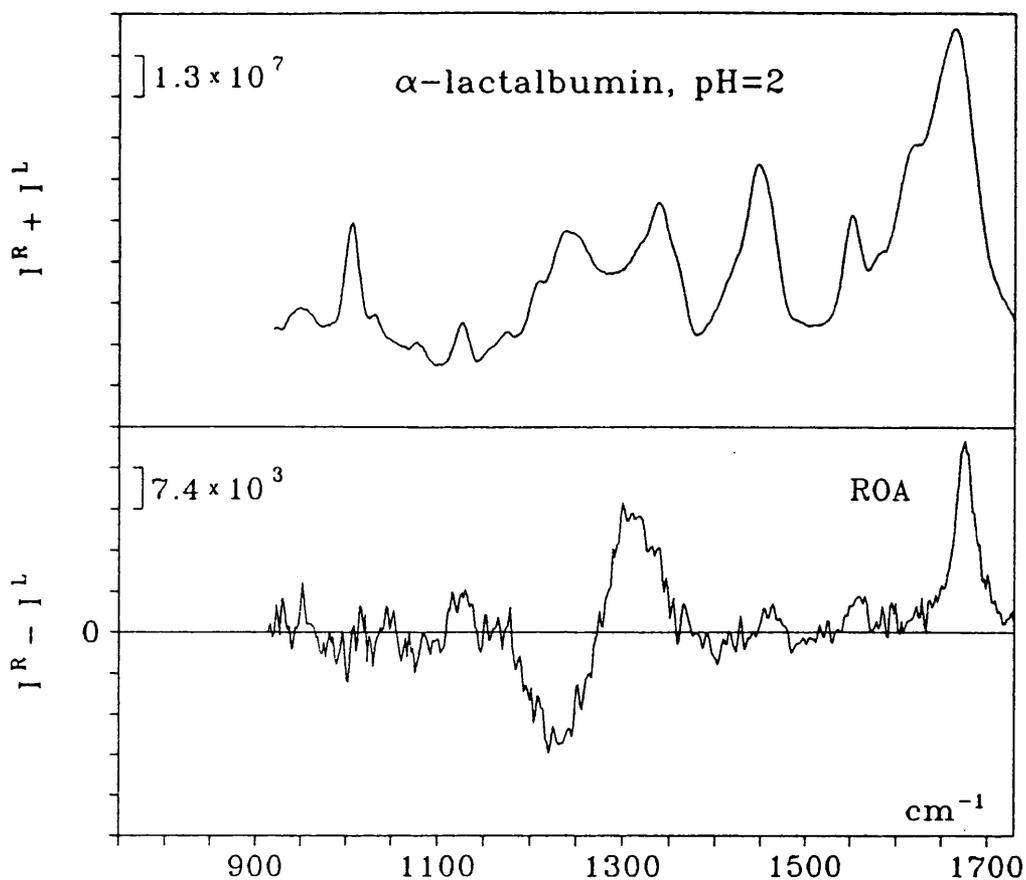


Fig. 5.11 Raman and ROA spectra of  $\alpha$ -lactalbumin (calcium bound) at pH=2 aqueous solution.

While the ROA spectra difference between the native and calcium free  $\alpha$ -LC are restricted to the resolution and intensity variation but the overall ROA pattern remains basically invariant, the adjustment of pH of native  $\alpha$ -LC solution from neutral to low pH results in drastic changes of the ROA spectra in the entire spectral region. Fig. 5.12 shows the comparison of the ROA spectra of the  $\alpha$ -lactalbumin in different forms and acidified state. The amide I ROA feature has altered from a negative-positive couplet in native  $\alpha$ -LC to a single sharp positive peak at

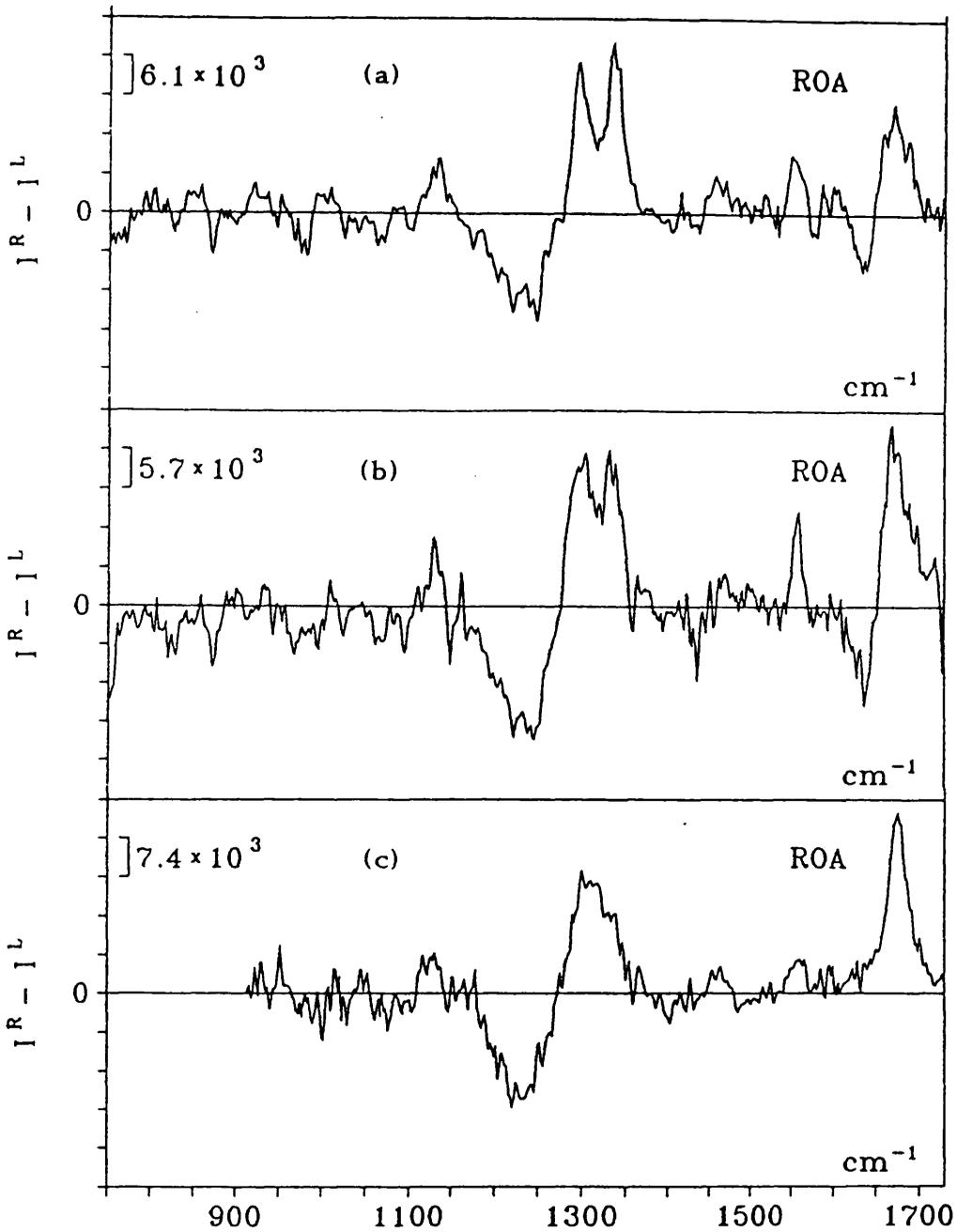


Fig. 5.12 A comparison of the three ROA spectra of  $\alpha$ -lactalbumin in different forms. (a), calcium bound, (b), calcium free, and (c), in low pH solution ( $\text{pH}=2$ ).

low pH and the amide III band now becomes a negative-positive conservative couplet centered at  $\sim 1270\text{cm}^{-1}$  negative at low frequency  $\sim 1240\text{cm}^{-1}(-)$  and positive at high  $\sim 1300\text{cm}^{-1}(+)$ . The ROA  $\Delta$  value is almost one order of magnitude different for the amide III bands between the two states. This is also in great contrast to the conventional Raman spectra between different pH value in which only minor changes in the amide III band region were noticed.<sup>22</sup>

The ROA spectra difference between the native and calcium free  $\alpha$ -LC on the one hand, prove the extreme sensitivity of ROA to structural order, and on the other indicates that there is considerable local structural change at the calcium binding site on removal of the calcium ion. The complete alteration of the ROA spectra upon acidification of  $\alpha$ -LC suggests that the polypeptide chain backbone probably is unfolded together with the loss of the calcium ions.

There is controversy on the issue of the conformation of the  $\alpha$ -LC in native and calcium free form. Early ECD and NMR study indicates that there is nearly identical ECD spectra between these two forms, and not much difference between the native and apo forms were observed.<sup>29,30</sup> On the contrary, recent report by White<sup>31</sup> shows that there is a clear difference in the ECD spectra in the 200-230nm region between the two forms of  $\alpha$ -LC.  $^1\text{H}$  NMR spectra also reveals that there are considerable changes upon calcium depletion.<sup>31</sup>

Other kinds of data also demonstrated that some structural changes take place upon removal of the calcium ion. The molecular volume was

increased 30% in calcium free  $\alpha$ -LC<sup>21</sup> which means the calcium free form is much looser than that with calcium bound. Permyakov suggested that there are conformational changes.<sup>33</sup> More recently, Rao et al. concluded that unfolding takes place upon acidification.<sup>34</sup> Our ROA data clearly show that there is considerable structural difference between calcium bound and free states. It is high likely that the microenvironment at the calcium binding site is changed when the calcium ion is depleted.

#### 5.4 Conclusions

The eight protein ROA spectra demonstrate convincingly that high quality ROA spectra can now be recorded thanks to the improvement of ROA instrument sensitivity, which was a dream never realized before although the potential of ROA for the study of proteins has long been recognized. The much wider spectral range accessible to ROA measurement contrasts sharply with the rather limited region amenable by VCD. One definite superiority of ROA over VCD is that ROA can be used to investigate proteins in aqueous solution both in water and D<sub>2</sub>O, whereas VCD is limited to D<sub>2</sub>O. As has been mentioned above, the use of D<sub>2</sub>O as solvent will inevitably replace all the labile protons in the amide group, which in turn will change the characteristic vibrational spectra and vibrational optical activity spectra.

Comparing ROA with the popular technique ECD, one can see that a ROA spectrum possesses an enormous information content involving every

part of the molecule while a ECD spectrum usually contains a few overlapped bands. On the other hand, ECD enjoys much greater sensitivity that enables it to probe protein conformation with very dilute solutions and protein dynamics with time-resolved mode. It can be predicted that ROA will become a powerful technique in protein conformation study in solution and will provide much more useful information with respect to the secondary structure and probably the tertiary structure.

Examining the ROA spectra of the eight proteins measured so far and the two model polypeptides reveals that ROA features of proteins are chiefly associated with molecular groups on the polypeptide backbone, e.g. C=O, N-H/C-H, C<sub>α</sub>-N, C<sub>α</sub>-C groups and that the side groups do not make significant contributions except for a few aromatic amino acid residues like tryptophan and tyrosine. This is particularly manifested in the ROA spectrum of BSA which contains 582 amino acid residues but no significant ROA features observed can be associated with the side groups.

The ROA spectra of the eight proteins shows that the most intriguing ROA features are concentrated in the extended amide III region where each protein exhibits its individual pattern. Although it is difficult to decipher these characteristic ROA pattern at the moment, it is conceivable that they contain rich information about the protein chain conformation that could be extracted in the future when ROA data on more proteins are accumulated and the correlation between the ROA pattern and structure becomes established.

It is important to note that the  $\Delta$ -value of protein ROA bands is a valuable source of information with respect to protein chain conformation apart from band frequency and sign. The  $\Delta$ -value of ROA bands at similar frequency could be several orders of magnitude different and it can change dramatically upon change of pH value. This indicates that the ROA  $\Delta$ -value is a very sensitive probe to structural changes. It is expected that, with further improvement of overall performance of the ROA instrument, the ROA  $\Delta$ -value could be evaluated more precisely in order to be able to compare them between different samples and provide useful information.

The dramatic ROA spectra variation upon pH change in  $\alpha$ -lactalbumin demonstrates that ROA is much more sensitive than conventional Raman spectra to the local structural change at ligand binding site. Thus, ROA offers a sensitive probe to the protein-ligand interaction investigations, which is particularly demanding.

It should be pointed out that the routine measurement of protein ROA spectra and the wide application of ROA spectroscopy in protein chemistry are still challenges for ROA spectroscopists. The principle barriers at present are the instrument sensitivity and fluorescence interference from sample and impurities. The limitation of the instrument sensitivity is expected to be removed with further instrument development by adopting a new collection optic and from advances in optical and electronic technologies. Fluorescence interference could be practically reduced by more sophisticated physico-chemical purification methods.

## REFERENCES

1. C. K. Mathews and K. E. van Holde, *Biochemistry*, The Benjamin/Cummings Publishing Company, Inc, New York, 1990.
2. T. E. Creighton, *Proteins: Structure and Molecular Properties*, Freeman, San Francisco, 1983.
3. P. R. Carey, *Biochemical Application of Raman and Resonance Raman Spectroscopies*, Academic Press, New York, 1982.
4. G. M. Glore and A. M. Gronenborn, *Protein Engineering*, **1**, 275, 1987.
5. B. G. Frushour and J. L. Koenig, *Adv. IR and Raman Spectrosc.*, edited by R. J. H. Clark and R. E. Hester, John Wiley & Sons, New York, **1**, 35, 1976.
6. T. G. Spiro, *Chemistry in Britain*, 602, 1989.
7. S. Krimm and J. Bandekar, *Biopolymers*, **19**, 1, 1980.
8. G. D. Rose, L. M. Gierasch and J. A. Smith, *Adv. Prot. Chem.*, **37**, 1, 1985.
9. S. Krimm, in '*Biological Application of Raman Spectroscopy*', Ed T. G. Spiro, **1**, 1, 1987.
10. P. Lagant, G. Vergoten, G. Fleury and M. H. Loucheux-Lefebvre, *Eur. J. Biochem.*, **139**, 137, 1984.
11. J. A. Fox, A. T. Tu, V. J. Hruby and H. J. Mosberg, *Arch. Biochem. Biophys.* **211**, 628, 1981.
12. R. W. Woody, in '*Peptides, Polypeptides and Proteins*', edited by E. R. Blout, F. A. Bovey, M. Goodman and N. Lotan, John Wiley and Sons, New York, 338, 1974.

13. J. S. Richardson, *Adv. Prot. Chem.*, **34**, 167, 1981.
14. T. Peters Jr, *Adv. Prot. Chem.*, **37**, 161, 1985.
15. T. L. Blundell et al, *Nature*, **231**, 506, 1972.
16. S. J. Ford et al. to be published
17. X. M. He and D. C. Carter, *Nature*, **358**, 209, 1992.
18. D. C. Phillips, *Proc. Natl. Acad. Sci., U.S.A.*, **57**, 484, 1967.
19. A. Fersht, *Enzyme Structure and Mechanism*, Freeman, San Francisco, 1977.
20. M. J. Kronman, *CRC Crit. Rev. Biochem.*, **24**, 565, 1989.
21. H. A. Mackenzie and F. H. White Jr, *Adv. Prot. Chem.*, **41**, 173, 1991.
22. N. T. Yu, *J. Am. Chem. Soc.*, **96**, 4664, 1974.
23. I. Harada, T. Miura and H. Takeuchi, *Spectrochim. Acta*, **42A**, 307, 1986.
24. L. Hecht, L. D. Barron, M. Lindner and B. Schrader, *11th International Conference on Raman Spectroscopy*, edited by R. J. H. Clark and D. A. Long, John Wiley & Sons, New York, 945, 1988.
25. L. Sawyer, M. Z. Pariz, A. C. T. North and E. E. Eliopoulos, *Bioch. Soc. Trans*, **13**, 265, 1985.
26. V. M. Naik and S. Krimm, *Int. J. Pept. Prot. Res.*, **23**, 1, 1984.
27. J. Prestrelski, D. M. Byler and M. P. Thompson, *Biochemistry*, **30**, 8797, 1991.
28. M. Urkanova, R. K. Dukor, P. Pancoska, V. P. Gupta and T. A. Keiderling, *Biochemistry*, **30**, 10479, 1991.
29. K. Kuwajima, Y. Hiraoka, M. Ikeyuchi and S. Sugai, *Biochemistry*, **24**, 874, 1985.
30. K. Kuwajima, Y. Harushima and S. Sugai, *Int. J. Pept. Prot. Res.*, **27**, 18,

1986.

31. F. H. White, *Internatl. J. Pept. Prot. Res.*, **39**, 265, 1992.
32. E. A. Permyakov, V. V. Yarmolenko, L. P. Kalinichenco, L. A. Morozova and E. A. Burstein, *Biochem. Biophys. Res. Commun.*, **100**, 191, 1981.
33. L. Lindahl and H. J. Vogel, *Anal. Biochem.*, **140**, 394, 1984.

## Raman optical activity of carbohydrates

### 6.1. Introduction

Carbohydrates represent one of the most important classes of biological molecules. However, the study of the structure and conformation of carbohydrates is well behind proteins and nucleic acids possibly owing to the lack of suitable techniques and the complexity of carbohydrate structure. There has been a resurgence in carbohydrate study in recent years and the carbohydrates have now been recognized as playing a crucial role in many bioactivities apart from its major role as energy storage and the structural material of the cell wall.<sup>1</sup>

Valuable information on carbohydrate conformation in aqueous solution ( $D_2O$ ) has been provided by N.M.R. spectroscopy.<sup>2</sup> However, N.M.R. is fundamentally not sensitive to chirality and only provides chiral information indirectly *via* chemical modification using a chiral auxiliary.<sup>3</sup> Furthermore, most carbohydrates are not favourable samples for conventional chiroptical studies using electronic circular dichroism (ECD) because they absorb below the short wavelength limit of 190 nm of most commercial ECD instruments.<sup>4</sup> Vacuum ultraviolet circular dichroism UVCD,<sup>5</sup> has to be employed.

Vibrational spectra, both infrared and Raman,<sup>6</sup> have been used to

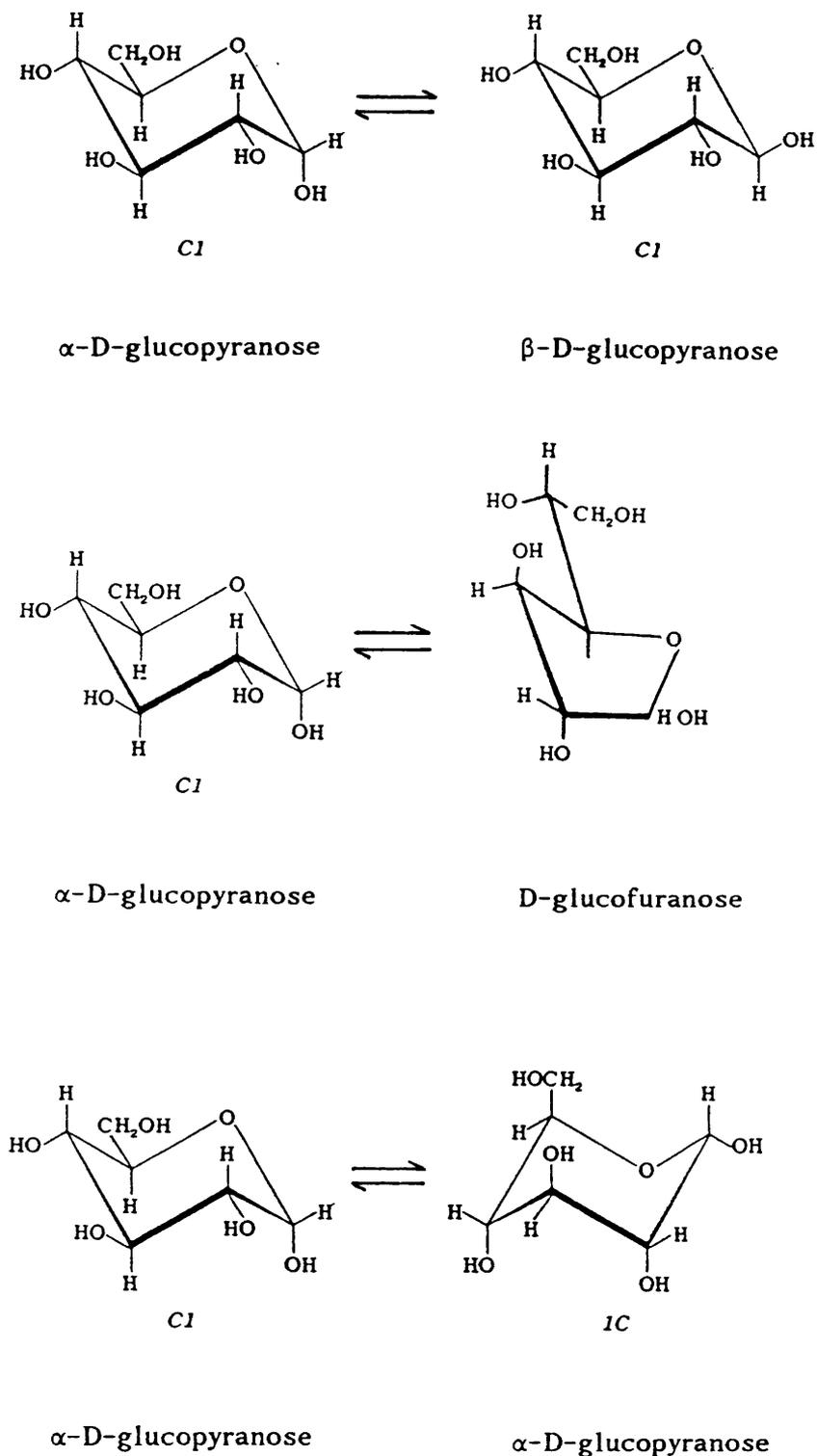


Fig. 6.1 The multiple equilibrium of D-glucose in aqueous solution.

investigate carbohydrate structure and have provided valuable information in solid phase. However, it is difficult to use infrared spectroscopy to study carbohydrates in aqueous solution because water absorbs strongly infrared radiation. Conventional Raman spectroscopy can be employed in carbohydrate studies in solution phase, but the usefulness of Raman approach to carbohydrates study have not been exploited as intensively as they should be probably because of the Raman spectra of carbohydrates are very complex and difficult to interpret.<sup>6</sup> The complexity of a carbohydrate structure is illustrated by Fig. 6.1. It shows a multiple equilibrium among different stereoisomers of D-glucose, which include ring tautomerism, anomerization and pyranoid ring conformation.

VCD has also been used to study carbohydrates including both the C-H stretch region<sup>7</sup> and the mid-infrared region.<sup>8</sup> Unfortunately, VCD faced problems associated with low instrument sensitivity and strong infrared absorption from water like its parent infrared spectroscopy though they demonstrated the potential value of vibrational optical activity studies in carbohydrate chemistry.

In contrast to ECD and VCD, carbohydrates are highly favourable for ROA studies.<sup>8</sup> The much more spectral information accessible to ROA as compared with ECD and VCD means that a characteristic 'chiral fingerprint' originating from the molecular architecture is immediately discernible. The much better resolution inherent in the optical activity property gives ROA a definite advantage over its parent Raman spectroscopy. In this chapter, we report detailed ROA results on a series of carbohydrate; namely, fifteen monosaccharides, one disaccharide and a

cyclodextrin to lay the foundations for the ROA studies of di-, oligo- and polysaccharide and their application in carbohydrates structure investigations.

## 6.2 Experimental section

The monosaccharides chosen comprise five hexapyranoses, namely D-glucose, D-galactose, D-mannose, D-allose and D-talose; five pentapyranoses, namely D-xylose, D-lyxose, D-arabinose, L-arabinose and D-ribose; and five methyl-pyranosides, namely methyl- $\alpha$ -D-glucoside, methyl- $\beta$ -D-glucoside, methyl- $\alpha$ -D-galactoside, methyl- $\beta$ -D-galactoside and methyl- $\alpha$ -D-mannoside. These monosaccharides differ from each other in the orientation of the hydroxyl groups, the dominant anomer configuration and the chair conformation in equilibrated aqueous solution. The disaccharide D-maltose and  $\alpha$ -D-cyclodextrin are chosen because they provide two good compounds to test the ROA capability to discriminate the glycosidic linkage which is of paramount important in polysaccharides. All these carbohydrates were purchased from Sigma except for D-glucose which was supplied by BDH.

All the samples were dissolved in distilled water and prepared as near-saturated solutions. This gave concentrations in the range 2M-5M depending on the solubility of the individual sugar at room temperature. For samples showing significant fluorescence background, filtration through charcoal was employed to remove the fluorescing impurities. The solutions were allowed to equilibrate overnight and filtered into quartz

microfluorescence cells using a 0.22 $\mu$  Millipore membrane filter (to avoid light scattering from dust particles) and centrifuged for half an hour before the ROA measurements.

All the ROA spectra were recorded with the following instrumental conditions: argon ion laser wavelength 514.5 nm, power at the sample 600 mW, spectrograph slit at 120 $\mu$  corresponding to a spectral bandwidth  $\sim 12$  cm<sup>-1</sup>, acquisition time 2 hours. The room temperature was kept to  $\sim 20^\circ$ C during the acquisition of the spectra.

Backscattered ROA spectra in the range 600-1600 cm<sup>-1</sup> were recorded: we could not record reliable ROA spectra below  $\sim 600$  cm<sup>-1</sup> in aqueous solution because of stray light fluctuations which become severe in the low frequency region.

### 6.3 Results and Discussion

The backscattered ROA spectra of these carbohydrates in the region between 600-1600cm<sup>-1</sup> are displayed in Figs. 6.2-6.15. It can be seen that most ROA bands of carbohydrates have  $\Delta$ -values in the range  $10^{-3} - 10^{-4}$  and that each sugar exhibits an individual ROA pattern together with some common characteristic bands. The richness of these ROA spectra suggests that they contain a wealth of information on the details of the sugar architecture.

The ROA spectra of monosaccharides can be conveniently divided into three regions. We call the region between  $\sim 750-950$  cm<sup>-1</sup> as the 'anomeric region' since we have found that ROA sign patterns in this range are

sensitive to different anomeric configurations of carbohydrates. Our nomenclature differs from that adopted in the conventional vibrational spectroscopy of carbohydrates<sup>6</sup> where this is called the 'fingerprint region'. We shall reserve the name 'fingerprint' for the region between 950-1200  $\text{cm}^{-1}$  where carbohydrates generally show characteristic bands associated with the sugar ring structure: this is a region where the vibrational bands of carbohydrates are overlapped and many ROA features appear; it is associated with many coupled vibrational modes in carbohydrates, such as C-O and C-C stretches and C-O-H deformations. In the region above  $\sim 1200 \text{ cm}^{-1}$ ,  $\text{CH}_2$  and C-O-H deformations both from the sugar ring and any exocyclic hydroxymethyl groups contribute significantly to the normal modes.

To derive correlations between ROA spectra and the detailed structure of carbohydrates is a challenging task. The complete theoretical analysis of ROA spectra from *ab initio* calculations<sup>9</sup> seems very promising for small organic molecules and has been successfully applied to molecules as large as alanine<sup>10</sup> and tartaric acid;<sup>11</sup> but unfortunately monosaccharides are just too large for *ab initio* ROA calculations at the present time. Simple models based on the bond polarizability theory have given insights into the generation of ROA signals from simple local groups although it may be difficult to extract information of general use.<sup>12, 13</sup> The most common approach practiced in conventional vibrational spectroscopy and also in ECD is the empirical correlation between the observed spectrum and particular molecular groups. With the accumulation of more ROA data on series of related molecules, many correlations between ROA spectra

and stereochemical details will become established.

### 6.3.1 Absolute configuration

Fig. 6.2 shows the ROA spectra of a pair of enantiomers, D-arabinose and L-arabinose. The good reflection symmetry of the two spectra demonstrates the reliability of ROA data of carbohydrates obtained on the instrument and clearly shows that ROA spectroscopy, being a chiroptical technique, can readily provide information on the absolute configuration of carbohydrates which is inaccessible using conventional vibrational spectroscopy.

### 6.3.2 Anomeric configuration

Next we consider two pairs of monosaccharides: methyl- $\alpha$ - and methyl- $\beta$ -D-glucoside, and methyl- $\alpha$ - and methyl- $\beta$ -D-galactoside. Each pair has opposite absolute configurations at the anomeric carbon atom but the same configuration at all the other asymmetric carbon atoms on the pyranoid ring, together with the same pyranoid ring conformation, and so should provide a good test of the sensitivity of ROA to anomeric configuration.

Fig. 6.3 and 6.4 showing the ROA spectra of methyl- $\alpha$ -D-glucoside and methyl- $\beta$ -D-glucoside reveal several major differences. Of particular significance is a strong positive ROA band at  $\sim 970 \text{ cm}^{-1}(+)$  present in methyl- $\beta$ -D-glucoside but not in methyl- $\alpha$ -D-glucoside which can be

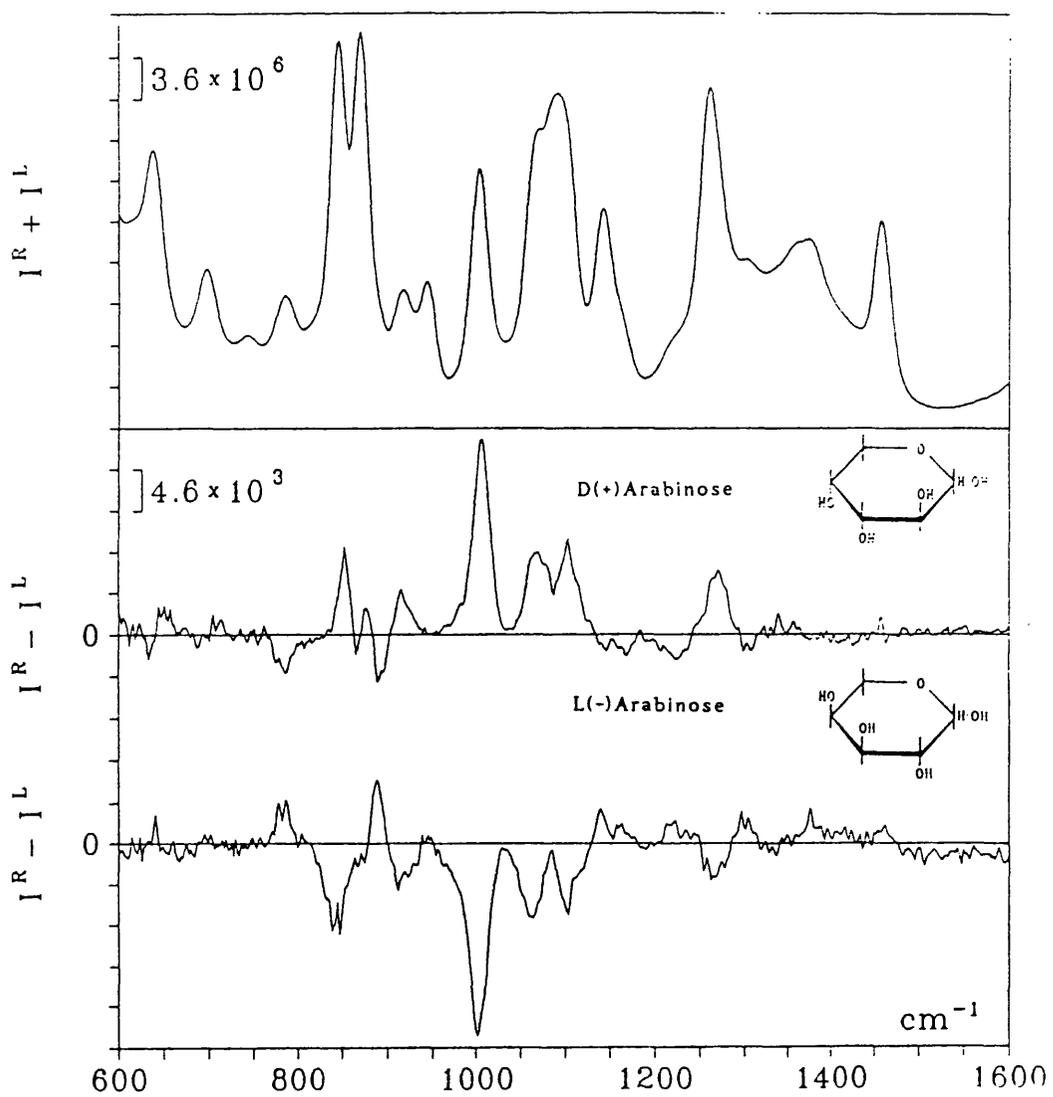


Fig. 6.2 Raman and ROA spectra of D-, L-arabinose in water.

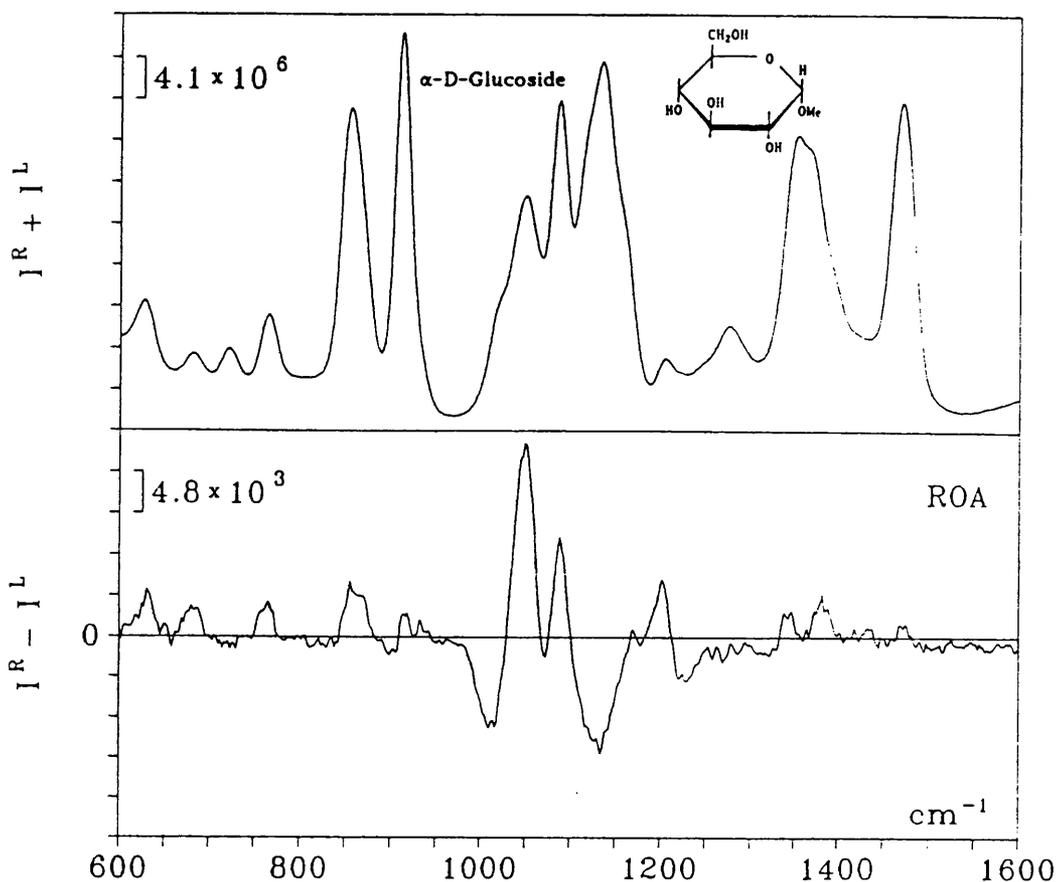


Fig. 6.3 Raman and ROA spectra of methyl- $\alpha$ -D-glucoside in water.

attributed to the different anomeric configurations. Since a conventional infrared study of methyl-pyranosides concluded that the C–O stretch mode of the methyl-ether group occurs in this frequency region,<sup>14,6</sup> this ROA band can probably be associated with the C–O stretch of the methoxyl group which is equatorial in methyl- $\beta$ -D-glucoside and axial in methyl- $\alpha$ -D-glucoside.

Other differences between the ROA spectra of the two anomers appear in the 1200–1500  $\text{cm}^{-1}$  region where methyl- $\beta$ -D-glucoside shows an ROA

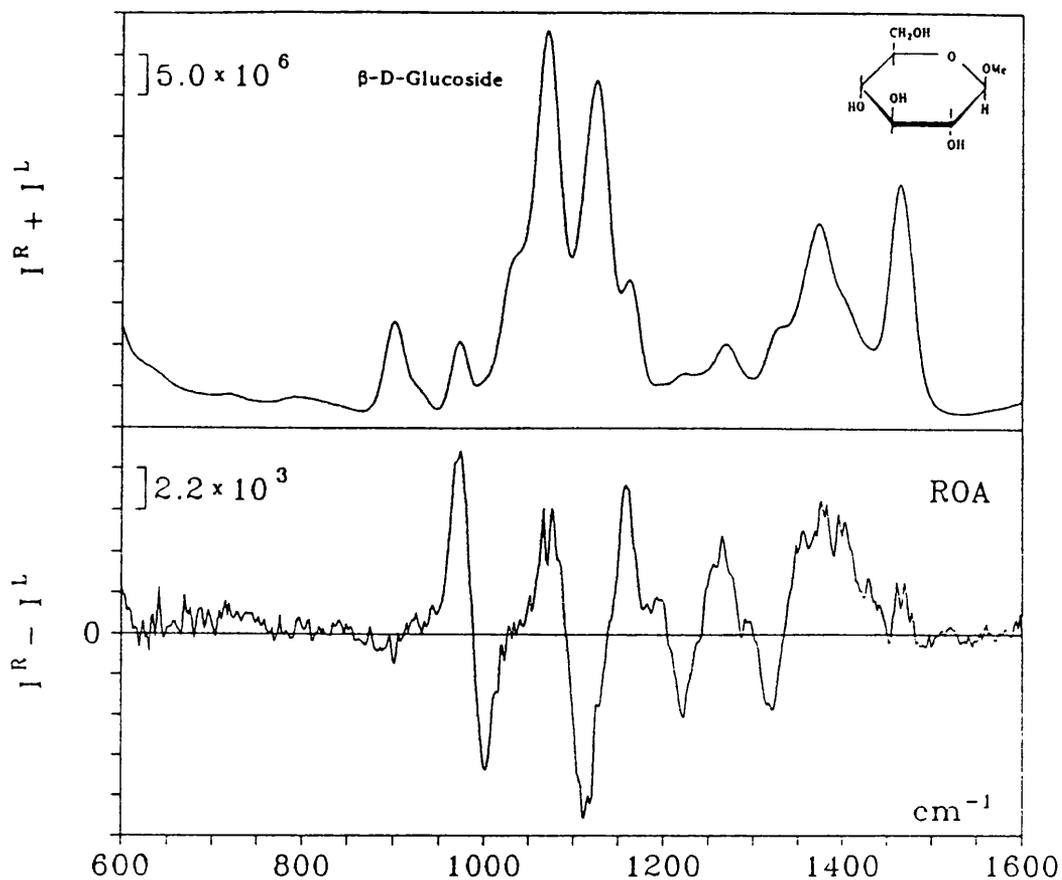


Fig. 6.4 Raman and ROA spectra of methyl- $\beta$ -D-glucoside.

couplet centered at  $\sim 1350 \text{ cm}^{-1}$ , negative at lower frequency and positive at higher, which disappears in the ROA spectrum of the methyl- $\alpha$ -D-glucoside. A similar ROA couplet is shown by D-glucose, but not in D-xylose, which suggests that it originates in vibrations of the hydroxymethyl group at C-5. This indicates that the local structure of the hydroxymethyl group at C-5 is influenced by the anomeric configuration, the appearance or otherwise of this couplet reflecting the different conformations of this substituent.

Striking differences are also observed between the ROA spectra of methyl- $\alpha$ -D-galactoside and methyl- $\beta$ -D-galactoside (Fig. 6.5 and 6.6). A strong ROA band at  $\sim 870\text{cm}^{-1}$  is particularly prominent which has similar magnitude but opposite sign for the two anomers. In fact the details of the two ROA spectra between  $\sim 800\text{--}950\text{ cm}^{-1}$  show quite close mirror symmetry, which implies that the vibrational modes responsible for these bands are in at least approximately enantiomorphous environments in the  $\alpha$  and  $\beta$  anomers although a proper assignment of these bands to specific vibrational modes is not available.

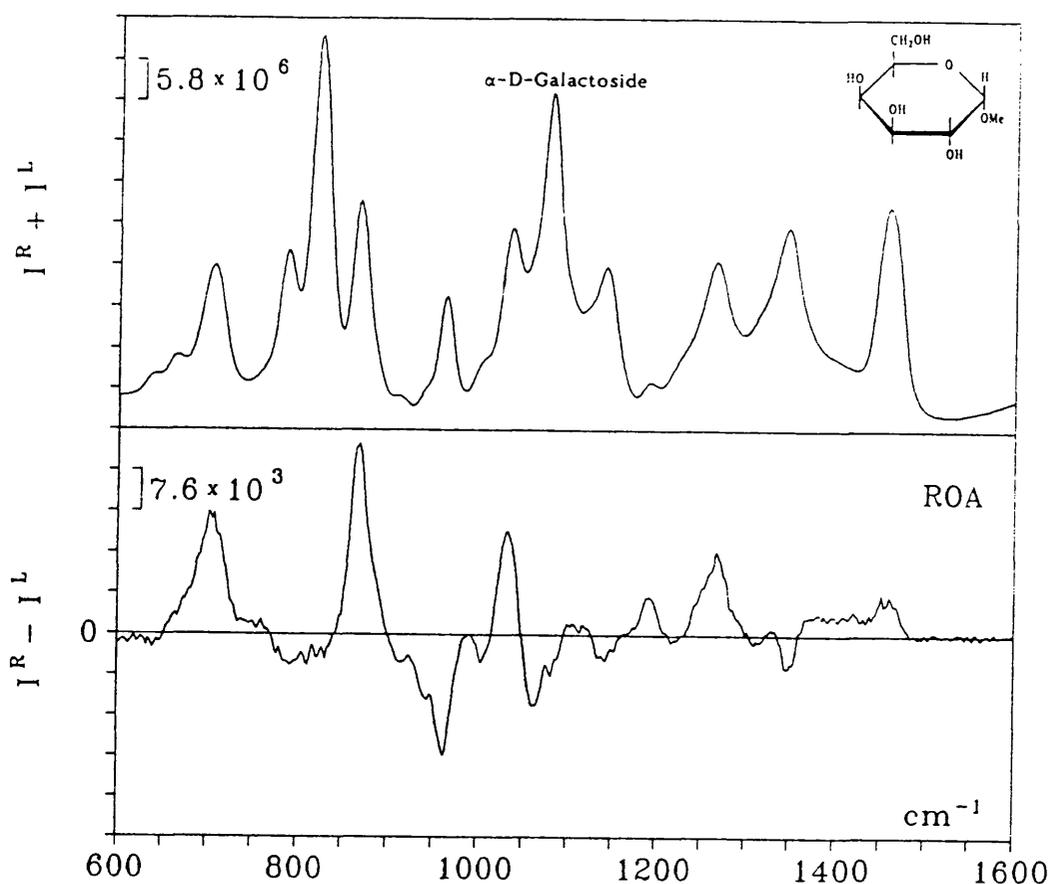


Fig. 6.5 Raman and ROA spectra of methyl- $\alpha$ -D-galactoside in water.

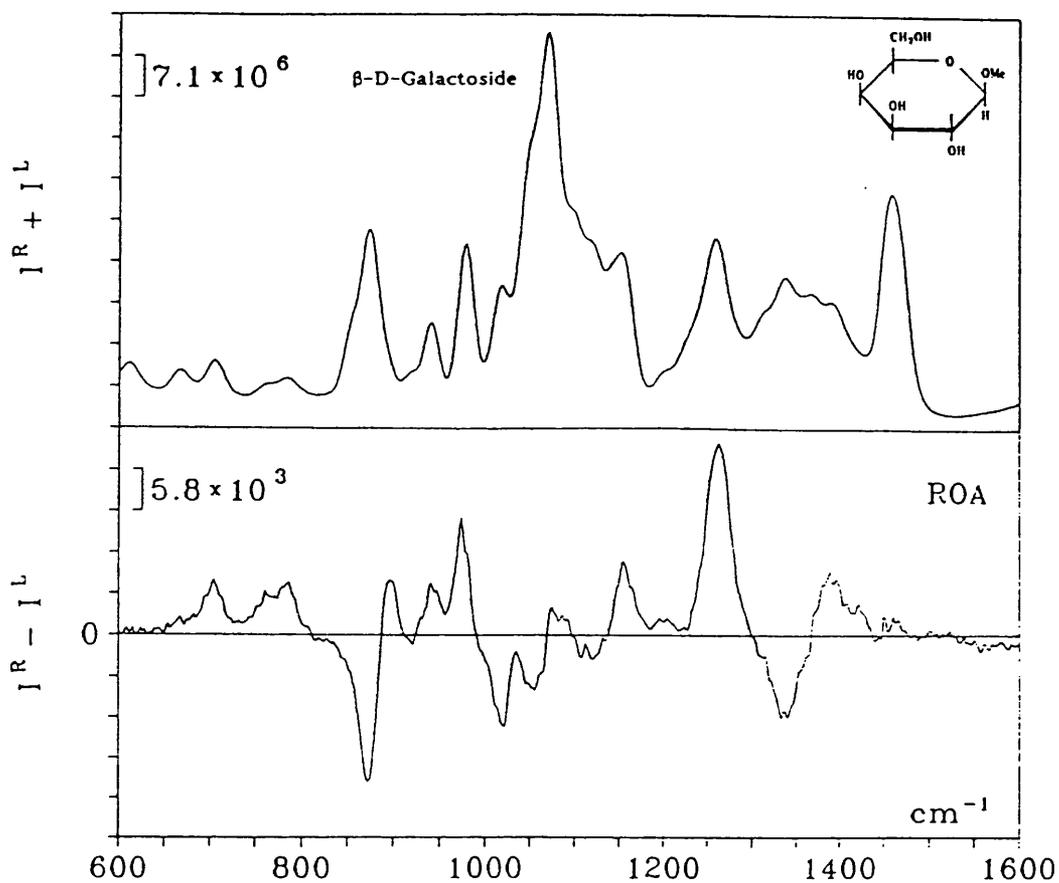


Fig. 6.6 Raman and ROA spectra of methyl- $\beta$ -D-galactoside in water.

Other differences are readily discernible elsewhere. Methyl- $\beta$ -D-galactoside shows a huge positive ROA band at  $\sim 1267 \text{ cm}^{-1}(+)$  while the magnitude of this band is considerably reduced in the methyl- $\alpha$ -D-galactoside. In contrast, another positive band at  $\sim 702 \text{ cm}^{-1}(+)$  increases its intensity from methyl- $\beta$ -D-galactoside to methyl- $\alpha$ -D-galactoside. As in the methyl-glucoside pair, some of the differences in the  $1200\text{--}1500 \text{ cm}^{-1}$  region also presumably reflect the different influence of the anomeric groups on the hydroxymethyl group at C-5 and its interaction with the adjacent axial hydroxyl group at C-4.

Examination of the ROA spectra of all fifteen monosaccharides reveals two particularly noticeable ROA bands in the anomeric region. For example, In D-galactose these are at  $\sim 822(-)$  and  $\sim 891(+)$   $\text{cm}^{-1}$  with ROA bands at  $825(-)$  and  $870(+)$   $\text{cm}^{-1}$  in its homomorphic sugar L-arabinose. D-mannose shows bands at  $\sim 827(+)$  and  $886(-)$   $\text{cm}^{-1}$ , methyl- $\alpha$ -D-mannoside at  $833(+)$  and  $880(-)$   $\text{cm}^{-1}$ , and D-lyxose at  $\sim 838(+)$  and  $884(-)$   $\text{cm}^{-1}$ . The ROA band frequencies and signs of fifteen monosaccharides in the anomeric region are listed in the right column of Table 6-1. The correlation is matched well to the absolute configuration of the anomeric carbon of the predominant anomer of at least ten monosaccharides. Exceptions are the glucose-type monosaccharides D-glucose, D-xylose and D-allose which show different frequencies; and D-tallose and D-ribose for which a significant proportion of molecules in solution exist in the furanose form due to ring tautomerism.

Our conclusions are consistent with conventional infrared work which has established anomer bands of carbohydrates in this region.<sup>6,14</sup> Two infrared absorption bands at  $844 \pm 7 \text{cm}^{-1}$  and  $890 \pm 7 \text{cm}^{-1}$  called type 2b and 2a have been correlated with the  $\beta$ - and  $\alpha$ -anomers, respectively. They are believed to originate from vibrational modes involving C-1-H deformations. Raman studies and normal vibrational mode calculations on D-glucose<sup>15,16</sup> suggest that the  $\beta$ -anomer band at  $\sim 890 \text{cm}^{-1}$  is associated with C-1-H,  $\text{CH}_2$  and C-O-H deformations. The  $\alpha$ -anomer band at  $\sim 840 \text{cm}^{-1}$  was also assigned to the coupling of C-1-H and  $\text{CH}_2$  deformations, but without a contribution from C-O-H deformations. It is interesting to note that the prominent ROA couplet of the glycosidic linkage of oligosaccharides is located in the anomeric region (*vide infra*).

Table 6-1. The predominant anomeric configurations, substituent dispositions and pyranose ring conformations, and the correlations between ROA bands and predominant anomers, of fifteen monosaccharides.

	pyranose anomer	OR <sup>b</sup> dispositions				pyranose ring conformation	ROA sign pattern	absolute configuration	
		C1	C2	C3	C4				
D-glucose	64%β	E	E	E	E	C1(D)	920(+)	—	R
D-galactose	64%β	E	E	E	A	C1(D)	891(+)	822(-)	R
L-arabinose		E	E	E	A	C1(L)↔1C(L)	870(+)	825(-)	R
D-xylose	64%β	E	E	E	E	C1(D)	895(+)	—	R
D-ribose	56%β	E	E	A	E	C1(D)	884(+)	850(-)	R
D-allose	70%β	E	E	A	E	C1(D)	915(+)	—	R
β-D-glucoside	100%β	E	E	A	E	C1(D)	970(+)	—	R
β-D-galactoside	100%β	E	E	E	A	C1(D)	890(+)	870(-)	R
α-D-glucoside	100%α	A	E	E	E	C1(D)	—	—	
α-D-galactoside	100%α	A	E	E	A	C1(D)	910(-)	870(+)	S
α-D-mannoside	100%α	A	A	E	E	C1(D)	880(-)	833(+)	S
D-mannose	67%α	A	A	E	E	C1(D)	886(-)	827(+)	S
D-lyxose	72%α	A	A	E	E	C1(D)↔1C(D)	884(-)	838(+)	S
D-arabinose	60%α	A	A	A	E	1C(D)↔C1(D)	870(-)	825(+)	S
D-talose	40%α	A	A	E	A	C1(D)	887(-)	850(+)	S

<sup>a</sup> Taken from N.M.R. data in reference 16.

<sup>b</sup> R = H or CH<sub>3</sub>; E = equatorial orientation; A = axial orientation.

### 6.3.3 Homomorphic sugars

Homomorphic sugars have the same configuration on each asymmetrically substituted carbon atom of the pyranoid ring but differ in the type of substituent attached to one carbon atom. There are six pairs of homomorphic sugars among the fifteen monosaccharides investigated. Thus, the three pairs D-xylose and D-glucose, L-arabinose and D-galactose, and D-lyxose and D-mannose differ from each other by an exocyclic hydroxymethyl group at C-5. Presumably, the anomerization of these sugars should not affect the parallel comparison between paired homomorphic sugars because the first two pairs of monosaccharides favour the  $\beta$ -anomer with the same fraction (64%), and the sugars of the last pair are predominantly in the  $\alpha$ -anomeric form.<sup>2</sup> The other three pairs D-glucose and methyl- $\beta$ -D-glucoside, D-galactose and methyl- $\beta$ -D-galactose, D-mannose and methyl- $\alpha$ -D-mannoside differ by a methoxyl group at C-1. Although the latter three pairs are not truly homomorphic sugars because of the anomerization of the pyranose forms, the predominant  $\beta$ -anomers in D-glucose, D-galactose and predominant  $\alpha$ -anomer in D-mannose in equilibrated aqueous solution should make them comparable with the corresponding methyl-pyranosides.

Comparison of the ROA spectra for each pair of homomorphic sugars reveals similar ROA patterns in every case, particularly in the fingerprint region ( Fig. 6.8, 3.7; 6.2, 6.9; 6.10, 6.11; 3.7, 6.4; 6.9, 6.6; and 6.11, 6.7). D-glucose shows four ROA bands at  $\sim 1124\text{ cm}^{-1}(-)$ ,  $1065\text{ cm}^{-1}(+)$ ,  $1025\text{ cm}^{-1}(+)$  and  $1000\text{ cm}^{-1}(-)$  in the fingerprint region, which constitute a

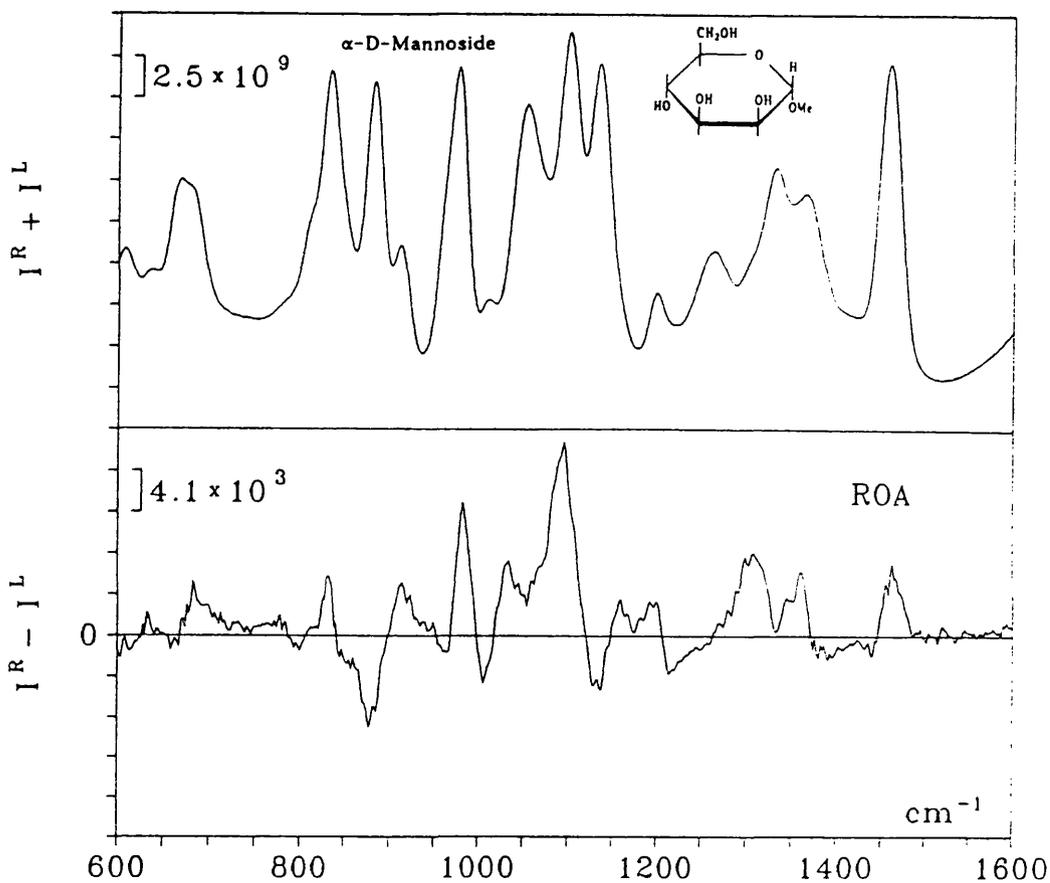


Fig. 6.7 Raman and ROA spectra of methyl- $\alpha$ -D-mannoside in water.

characteristic ROA pattern. D-xylose exhibits almost the same ROA pattern as that of D-glucose in the fingerprint region but differs in the region between  $\sim 1200$ - $1500 \text{ cm}^{-1}$  on account of the lack of a  $\text{CH}_2\text{OH}$  substituent at C-5. The ROA spectrum of methyl- $\beta$ -D-glucoside (Fig. 6.4) between  $1000$ - $1500 \text{ cm}^{-1}$  is almost identical to that of D-glucose except for a small ROA feature at  $\sim 1460 \text{ cm}^{-1}$  which is probably associated with the methyl group.

The similarity of the ROA spectra for the homomorphic pair D-galactose and L-arabinose can be seen from Fig. 6.9 and 6.2. D-galactose

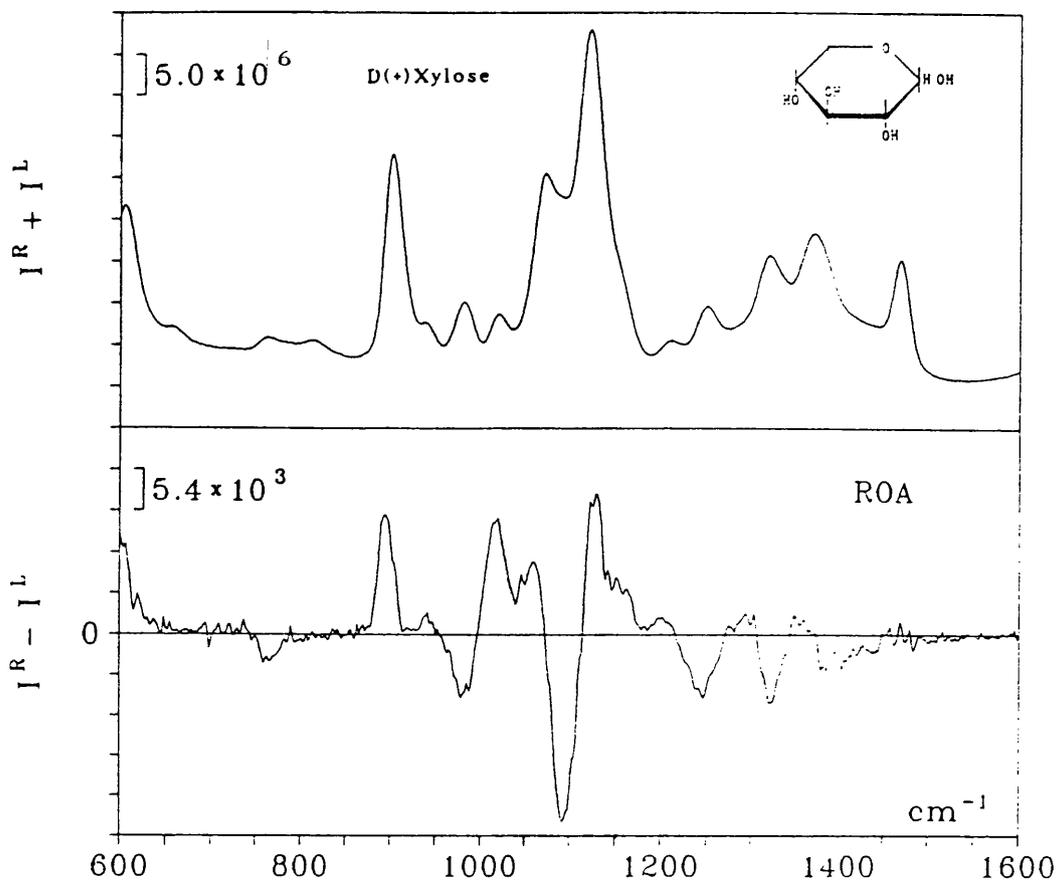


Fig. 6.8 Raman and ROA spectra of D-xylose in water.

shows three ROA bands at  $\sim 1149 \text{ cm}^{-1}(+)$ ,  $1065 \text{ cm}^{-1}(-)$ ,  $1020 \text{ cm}^{-1}(-)$  in the fingerprint region, a huge positive band at  $\sim 1267 \text{ cm}^{-1}(+)$  and a negative band at  $\sim 1341 \text{ cm}^{-1}(-)$  between  $1200\text{--}1500 \text{ cm}^{-1}$ ; while L-arabinose shows three similar ROA bands in the fingerprint region but different features in the  $1200\text{--}1500 \text{ cm}^{-1}$  region which can again be attributed to the lack of a  $\text{CH}_2\text{OH}$  group in L-arabinose. As mentioned before, their ROA features are almost the same in the anomeric region. The ROA spectrum of methyl- $\beta$ -D-galactose is almost identical to that of D-galactose between  $1050\text{--}1500 \text{ cm}^{-1}$ .

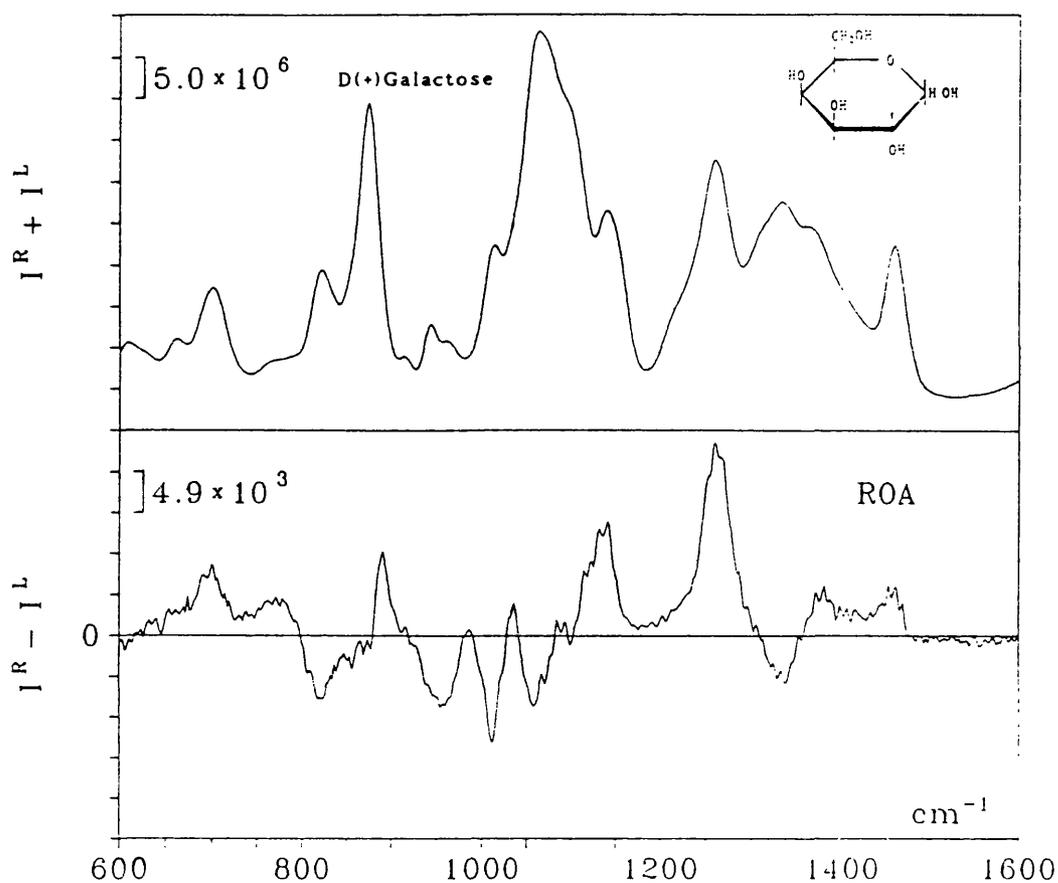


Fig. 6.9 Raman and ROA spectra of D-galactose in water.

Comparison can be made between another homomorphic pair of sugars, D-mannose and methyl- $\alpha$ -D-mannoside ( Fig. 6.10 and 6.7 ). They resemble each other between 1000-1500  $\text{cm}^{-1}$  except for a positive band at  $\sim 1460 \text{ cm}^{-1}$  in methyl- $\alpha$ -D-mannoside which again is related to the methyl group.

The vibrational bands of monosaccharides in the fingerprint region are mainly associated with C-O and C-C stretches and coupled C-O-H and C-C-H deformations.<sup>6</sup> It is believed that they are characteristic of the sugar ring structure although a detailed assignment of these vibrational

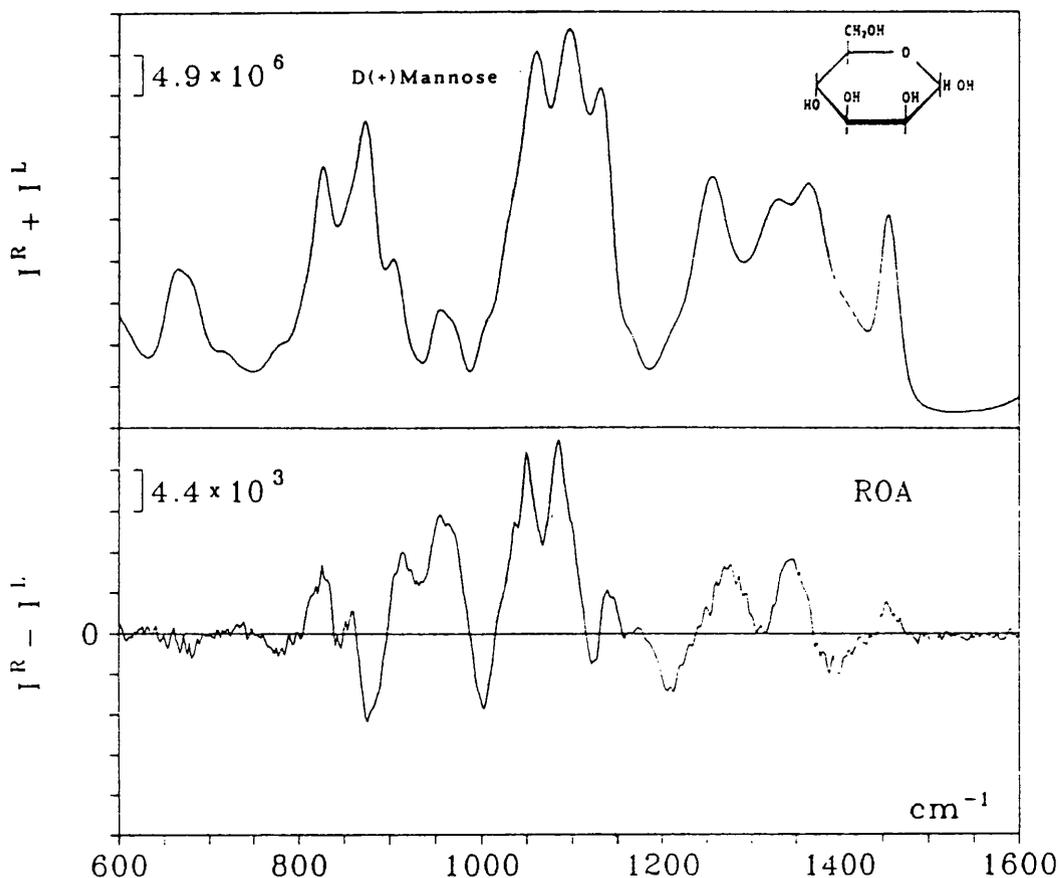


Fig. 6.10 Raman and ROA spectra of D-mannose in water.

bands is difficult due to mixing of endocyclic and exocyclic C–O stretches. Nevertheless, deuteration studies<sup>17</sup> on D-glucose, D-maltose, D-cellobiose and dextrans revealed that two bands at  $1070 \text{ cm}^{-1}$  and  $1020 \text{ cm}^{-1}$  originate in vibrations of the C–O–H group. A VCD study<sup>8</sup> of monosaccharides also suggested that VCD bands between  $1200\text{--}1100 \text{ cm}^{-1}$  are associated with the configuration of the C–O groups, and a band at  $1150 \text{ cm}^{-1}$  assigned to delocalized C–O stretches was thought to be influenced by the orientation of the exocyclic C–O and C–C groups.

The close similarity of the ROA spectra for each pair of homomorphic

monosaccharides, and the vast difference between epimeric sugars (*vide infra*) indicates that the ROA patterns, particularly in the fingerprint region, reflect the configuration of the substituents on the pyranoid ring (the conformation of which remains the same among homomorphic sugars), and that the introduction of a substituent on one exocyclic local group does not generally alter the ring conformation. These conclusions are supported by the fact that certain disaccharides and oligosaccharides show the basic ROA patterns of their monomer plus some additional ROA features originating in the linkages. Thus, the D-glucose-containing disaccharides D-maltose and D-cellobiose exhibit very similar ROA patterns to D-glucose in the fingerprint and 1200-1500  $\text{cm}^{-1}$  regions; similarly for cyclodextrins (*vide infra*). Also the disaccharides lactose and melibiose which contain a D-galactose unit show similar ROA spectra to D-galactose in the fingerprint region and the same huge positive ROA band at  $\sim 1267 \text{ cm}^{-1(+)}$ .<sup>18</sup>

The distinct ROA patterns of the three pyranoses D-glucose, D-galactose and D-mannose may well represent three typical substituent configurations on the pyranoid ring and should be of significant value for discriminating between the three most important classes of polysaccharides, namely the glucans, the galactans and the mannans.

The ROA spectra of the pair of homomorphic sugars D-lyxose and D-mannose, in contrast to the other five pairs of homomorphic sugars, are dramatically different from each other, particularly in the fingerprint region ( Figs. 6.10 and 6.11 ). D-mannose shows four ROA bands at  $\sim 1134 \text{ cm}^{-1(-)}$ ,  $\sim 1093 \text{ cm}^{-1(+)}$ ,  $\sim 1060 \text{ cm}^{-1(+)}$  and  $\sim 1012 \text{ cm}^{-1(-)}$  while D-lyxose

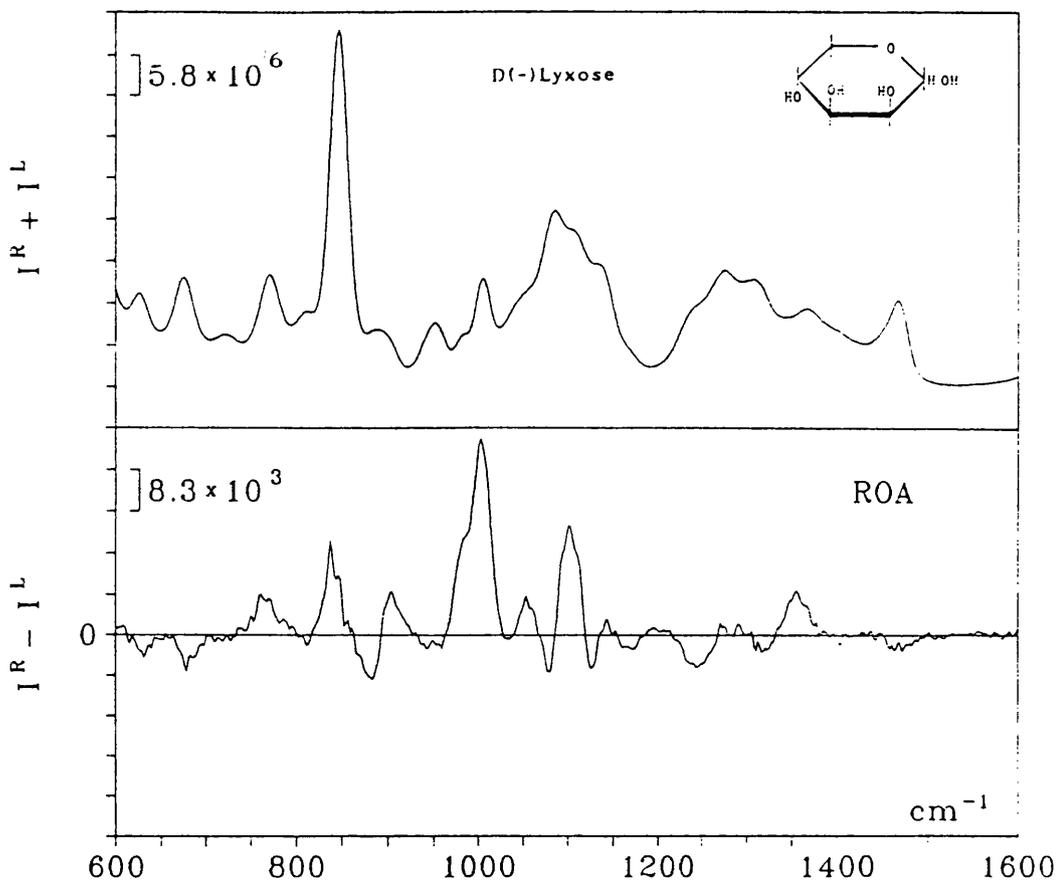


Fig. 6.11 Raman and ROA spectra of D-lyxose in water.

presents a completely different ROA pattern with a strong positive band at  $\sim 1005 \text{ cm}^{-1}$  which is not shown by D-mannose. The  $\Delta$ -value of this ROA band is of the order of  $10^{-3}$ , which is an order of magnitude larger than typical  $\Delta$ -values in monosaccharides. The origin of these significant ROA differences between a pair of homomorphic monosaccharides must be sought in structural factors other than anomeric configuration and OH dispositions because they are virtually the same. Lack of the exocyclic hydroxymethyl group at C-5 in D-lyxose is unlikely to be responsible for the differences in the fingerprint region. One possible explanation is that

they might be due to a different pyranoid ring chair conformation: as is well known from NMR studies,<sup>2</sup> D-lyxose exists as a mixture of C1(D) and 1C(D) conformations in equilibrated aqueous solution, but D-mannose, like all the other homomorphic pairs of sugars investigated, exists predominantly in the C1(D) chair conformation. This suggestion is reinforced by the fact that D-arabinose, a pentapyranose, has the same anomeric configuration as D-lyxose and also exists in two chair conformations in solution and shows a similar overall ROA spectrum. In particular, it also shows strong positive ROA at  $\sim 1005 \text{ cm}^{-1}$ . This close resemblance between the ROA spectra of the two pentapyranoses is probably not coincidental: all the other monosaccharides exist predominantly in the C1(D) conformation and none shows such a prominent ROA band at  $\sim 1005 \text{ cm}^{-1}$ .

#### 6.3.4 Epimeric sugars

Epimeric monosaccharides differ from each other by a different orientation of one substituent on the pyranoid ring. In contrast with homomorphic pairs of monosaccharides which show a considerable resemblance in the fingerprint and  $1200\text{--}1500 \text{ cm}^{-1}$  regions, epimeric pairs of sugars exhibit utterly different ROA features over the entire spectrum.

A comparison of epimeric monosaccharides which differ from each other by the disposition of a substituent at C-4 on the pyranoid ring can be made between D-glucose and D-galactose, methyl- $\alpha$ -D-glucoside and methyl- $\alpha$ -D-galactoside, and methyl- $\beta$ -D-glucoside and

methyl- $\beta$ -D-galactoside. Remarkable differences are observed in the ROA spectra between epimeric pairs in every case. Of most significance are the fingerprint and 1200–1500  $\text{cm}^{-1}$  regions since the ROA spectra of homomorphic pairs are usually very similar here. As mentioned in the previous section, D-glucose and D-galactose exhibit dramatically different ROA patterns in the fingerprint region ( Figs. 3.7 and 6.8 ). In the 1200–1500  $\text{cm}^{-1}$  region, D-glucose shows two couplets centered at  $\sim 1238 \text{ cm}^{-1}$  and  $1345 \text{ cm}^{-1}$ , both negative at lower and positive at higher frequency. In contrast, D-galactose gives a huge positive ROA band at  $\sim 1267 \text{ cm}^{-1}$  and a smaller negative band at  $\sim 1341 \text{ cm}^{-1}$ . Similar differences can be seen between methyl- $\beta$ -D-glucoside and methyl- $\beta$ -D-galactoside. These ROA differences apparently reflect the different dispositions of the C-4 hydroxyl group (which is equatorial in D-glucose and axial in D-galactose) and its interaction with neighbouring groups.

Comparing the ROA spectra of the epimers D-xylose and L-arabinose can expose the influence of the OH orientation at C-4. The ROA patterns are very different in the fingerprint region; and at higher frequency D-xylose shows two small negative bands at  $\sim 1245 \text{ cm}^{-1}$  and  $1326 \text{ cm}^{-1}$  while L-arabinose only presents a negative band at  $\sim 1265 \text{ cm}^{-1}$ . As stated before, the ROA pattern of a monosaccharide resembles more that of its homomorph than that of its epimorph. It is interesting to note that L-arabinose does not show a huge positive ROA band at  $\sim 1267 \text{ cm}^{-1}$  like that of D-galactose, which indicates that this band in D-galactose originates in intramolecular interference between the axial OH at C-4 and the hydroxymethyl group at C-5.

Comparing the ROA spectra of D-mannose and D-glucose, and of methyl- $\alpha$ -D-glucoside and methyl- $\alpha$ -D-mannoside, can expose the ROA signature of C-2 epimerization. As would now be expected, there is little resemblance between the ROA spectra of each pair of C-2 epimers. D-mannose displays four ROA bands in the fingerprint region which constitute a typical ROA pattern for carbohydrates of the D-mannose type. In the 1200-1500  $\text{cm}^{-1}$  region it shows two ROA couplets: the one centred at  $\sim 1250 \text{ cm}^{-1}$ , negative on the lower frequency side and positive on the higher, is very similar to that of D-glucose; the other centred at  $\sim 1375 \text{ cm}^{-1}$  is positive at lower frequency and negative at higher.

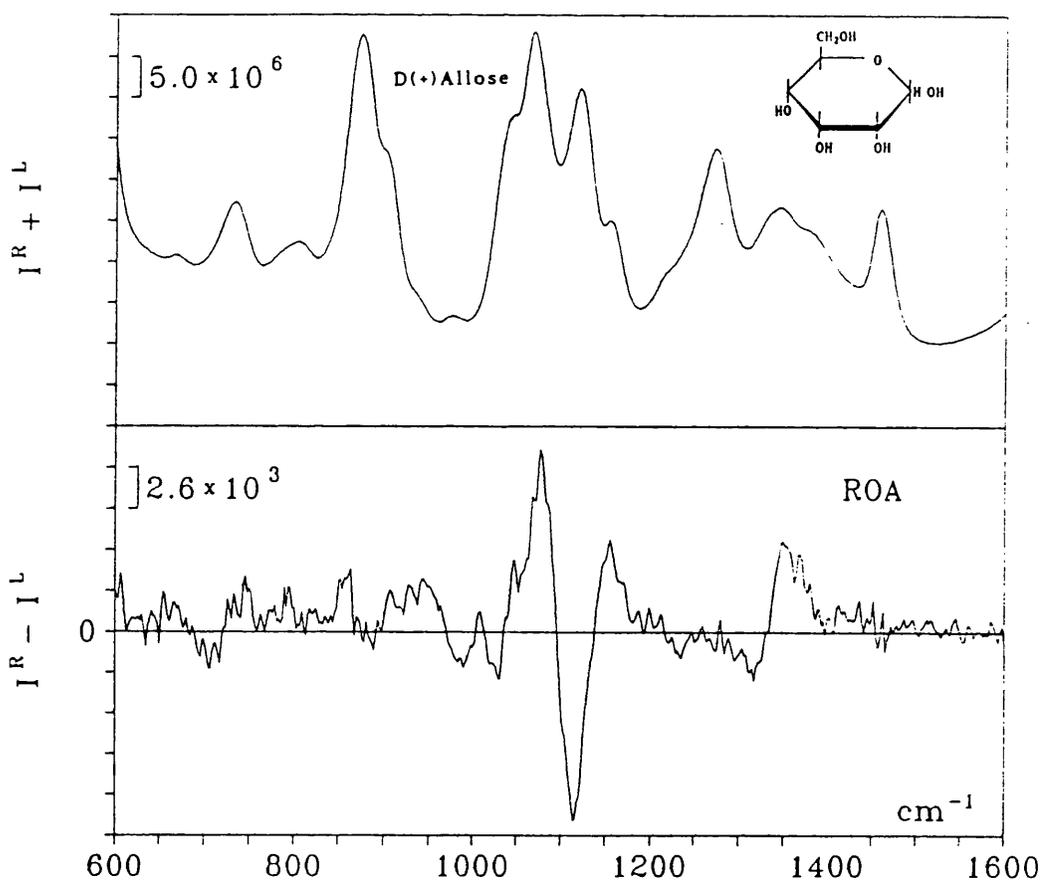


Fig. 6.12 Raman and ROA spectra of D-allose in water.

At first sight it is surprising that, unlike C-2 epimers, the ROA spectra of C-3 epimers are very similar. Thus the ROA spectrum of D-allose ( Fig. 6.12 ) is similar to that of D-glucose: they differ from each other by the configuration of a hydroxyl group at C-3 and both adopt predominantly the  $\beta$ -anomeric form in aqueous solution. Likewise the ROA spectra of D-lyxose and D-arabinose are similar, both sugars favouring the  $\alpha$ -anomeric form and also existing as a conformational equilibrium between C1(D) and 1C(D). These observations indicate that C-3 is a rather special position on the pyranoid ring since changing the disposition of an OH group here affects the ROA spectrum much less than at other positions. It is likely that there is little interaction of the OH group at C-3 with the hydroxymethyl group at C-5 or with the anomeric groups, whereas its counterpart at C-4 and C-2 interacts strongly with bulky adjacent groups at C-5 and C-1, respectively. This also suggests that the next-nearest interaction is not a significant factor in the generation of ROA spectra so that ROA probes very local stereochemical features.

### 6.3.5 Conformation of the exocyclic $\text{CH}_2\text{OH}$

The conformation of the exocyclic hydroxymethyl groups in monosaccharides has been intensively debated.<sup>19-21</sup> Knowledge of the preferred conformation of this group in polysaccharides is very important because it is associated with intramolecular hydrogen bonding and the stabilization of various chain backbone conformations.

It has been suggested that the hydroxymethyl group can exist in three

basic ideally staggered conformations, namely trans-gauche (TG), gauche-trans (GT) and gauche-gauche (GG) <sup>19</sup>. It was proposed that D-glucose prefers a gauche-trans conformation; and two possible conformations, gauche-trans (GT) and trans-gauche (TG), were suggested for D-galactose. D-mannose was considered to prefer a conformation differing from the other two. From NMR studies,<sup>20</sup> it was concluded that the  $\beta$ -anomer of D-galactose favors an 80% population of the trans-gauche conformation while the  $\alpha$ -anomer comprises equal amounts of the trans-gauche and gauche-trans conformers. The gauche-trans conformation is excluded on steric grounds for D-mannose.<sup>20</sup> Normal mode calculations<sup>21</sup> suggest that the most likely conformation of the CH<sub>2</sub>OH group is gauche-trans or gauche-gauche for  $\alpha$ -D-glucose and  $\alpha$ -D-mannose, and gauche-trans and trans-gauche for  $\alpha$ -D-galactose.

Raman bands in the 1200-1500 cm<sup>-1</sup> region are believed to originate mainly in deformations of the CH<sub>2</sub>, COH and the CH<sub>2</sub>OH groups.<sup>6</sup> In particular, two Raman bands at  $\sim 1263$  cm<sup>-1</sup> and  $\sim 1335$  cm<sup>-1</sup> were assigned as complex modes of the CH<sub>2</sub>OH group, the former from deuteration studies to a C-O-H bending mode and the latter to coupling of C-O-H bending and CH<sub>2</sub>OH twisting motions. Both bands are found to be conformationally sensitive in polysaccharides.<sup>15</sup>

The dramatic difference between the ROA patterns in the 1200-1500 cm<sup>-1</sup> region for D-glucose, D-galactose and D-mannose leads us to suggest that these ROA features reflect the different conformations or conformer population distributions of the exocyclic CH<sub>2</sub>OH group at C-5 which are expected to be strongly influenced by both the anomeric groups

and the disposition of the hydroxyl group at C-4. The conformations are basically the same between homomorphic sugars such as pyranoses and methyl-pyranosides as evidenced by the almost identical ROA spectra shown by D-glucose and methyl- $\beta$ -D-glucoside, D-galactose and methyl- $\beta$ -D-galactoside, and D-mannose and methyl- $\alpha$ -D-mannoside. The conformation of the CH<sub>2</sub>OH group in other monosaccharides can also be predicted from their ROA pattern in this region: thus, D-allose would have the same conformation as that of D-glucose; D-talose and D-ribose are similar to D-galactose; and methyl- $\alpha$ -D-glucoside and methyl- $\alpha$ -D-galactoside might exist as a combination of various conformations which could explain the cancellation of ROA features in this region. It should be pointed out that, although the exact conformation of the CH<sub>2</sub>OH group cannot be determined at present from ROA spectra, ROA does have the sensitivity to distinguish the different influence of anomeric groups and of the disposition of the hydroxyl groups at C-4 on the conformation of the CH<sub>2</sub>OH group. NMR measurement on  $\alpha$ - and  $\beta$ -galactosides shewed that the conformation of the exocyclic CH<sub>2</sub>OH group is influenced by the anomeric configuration.<sup>22</sup>

### 6.3.6 Furanose and pyranose

The last pair of monosaccharides we would like to compare is D-ribose and D-talose. Although the quality of their ROA spectra (Figs. 6.13 and 6.14) is not as good as those of the other monosaccharides, the main ROA features are clearly revealed. Despite the additional complexity produced by the ring tautomeric equilibrium, the two ROA spectra are very

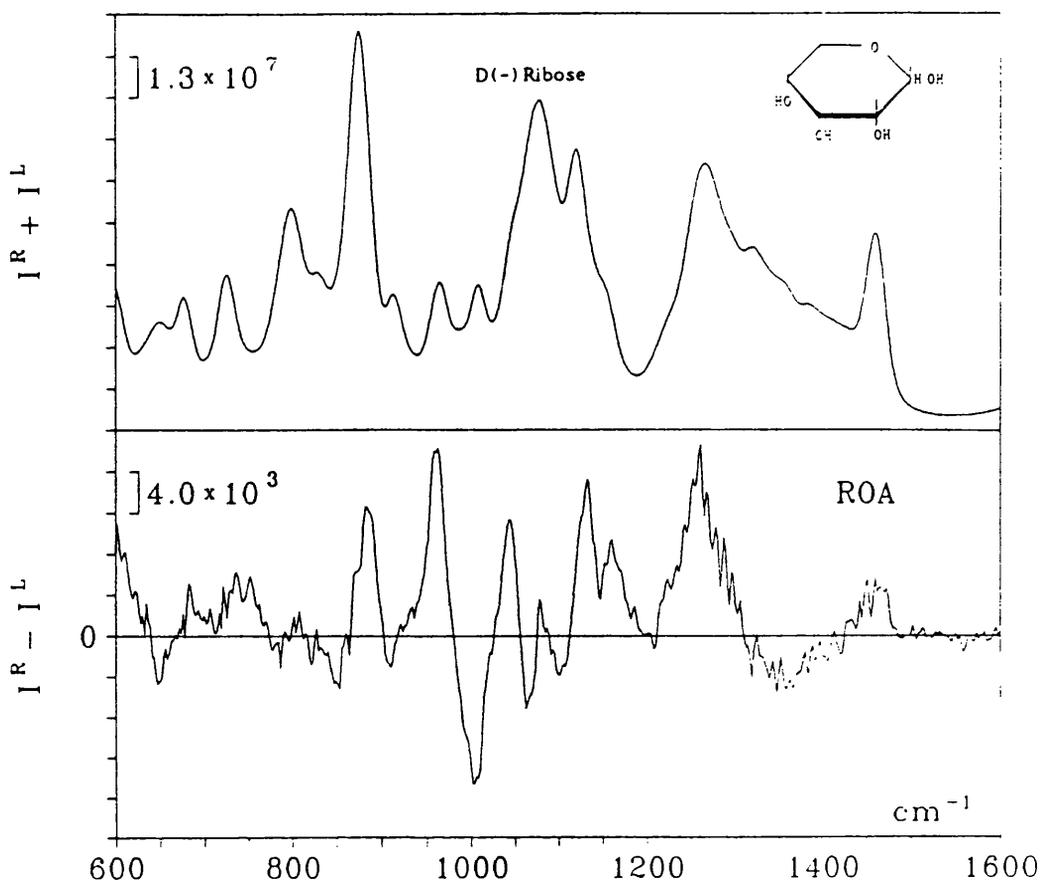


Fig. 6.13 Raman and ROA spectra of D-ribose in water.

similar in the 1200–1500  $\text{cm}^{-1}$  region, both showing a strong positive band at  $\sim 1250 \text{ cm}^{-1}$ (+) and a broad negative peak between 1320 and 1420  $\text{cm}^{-1}$ , which indicates that the  $\text{CH}_2\text{OH}$  group is in a similar environment. It may represent a characteristic ROA pattern of furanose since another monosaccharide D-fructose (not shown), which exists predominantly in the furanose form, also gives rise to a strong positive ROA band at  $\sim 1250 \text{ cm}^{-1}$ . The disaccharide sucrose which contains a D-glucose and a D-fructose residue also shows a positive ROA band at the same frequency but the intensity is considerably reduced due to the change of the local environment of the  $\text{CH}_2\text{OH}$  group on the furanose ring.<sup>18</sup>

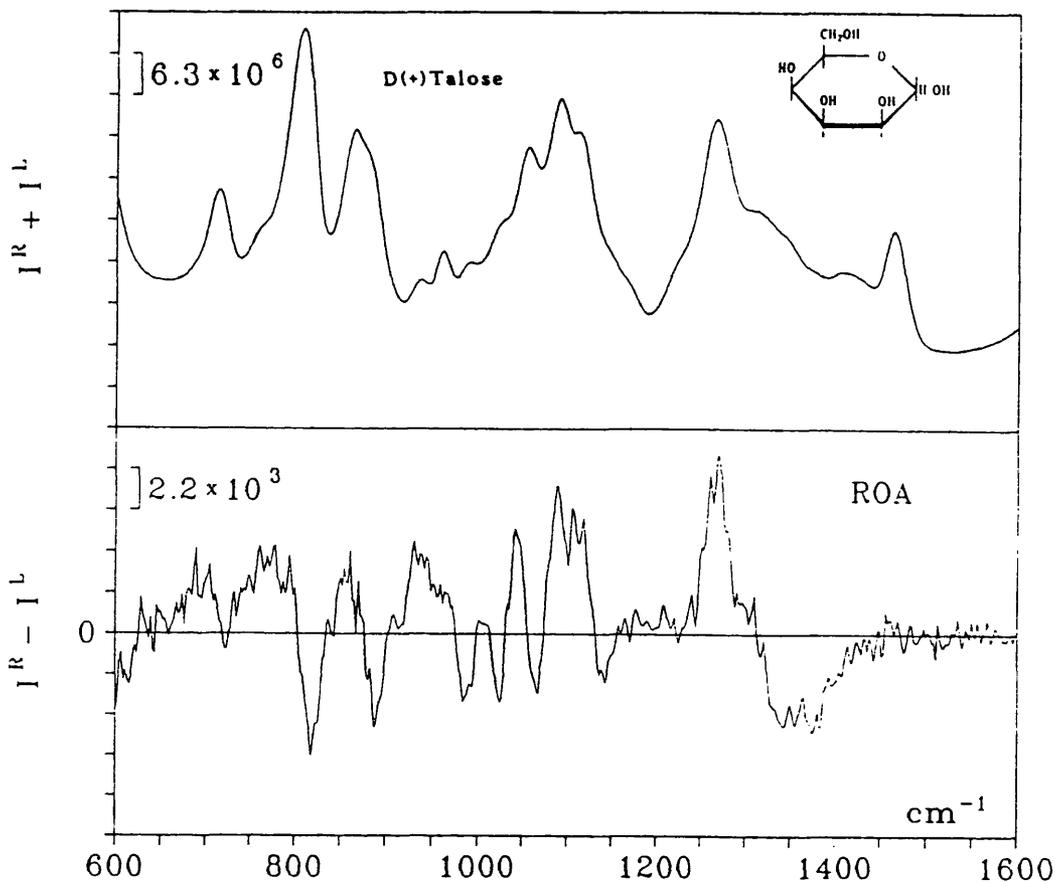


Fig. 6.14 Raman and ROA spectra of D-talose in water.

It is gratifying that the anomeric bands at  $\sim 884 \text{ cm}^{-1}(+)$  in D-ribose and  $887 \text{ cm}^{-1}(-)$  in D-talose manifest themselves well despite the ring tautomerism. This again confirms that ROA is extremely sensitive to the nearest group interactions.

It is worth pointing out that, as conventional Raman and FTIR studies of D-ribose and 2-deoxy-D-ribose suggest,<sup>23,24</sup> vibrational bands below  $700 \text{ cm}^{-1}$  are characteristic of the ribose ring and mainly associated with C-C-C and C-C-O deformations in the ring. A negative ROA band at

650  $\text{cm}^{-1}$  in D-ribose, 610  $\text{cm}^{-1}$  in D-talose, 615  $\text{cm}^{-1}$  in D-fructose and 650  $\text{cm}^{-1}$  in sucrose<sup>18</sup> can probably be attributed to the furanose form because none of the other monosaccharides shows such a feature in this region. We therefore anticipate that it might be possible to distinguish between pyranose and furanose rings from ROA spectra below  $\sim 700 \text{ cm}^{-1}$ , but detailed studies will have to wait until instrumental developments enable reliable ROA spectra of carbohydrates to be recorded in the low-frequency region.

### 6.3.7 The glycosidic linkage and the merit of the ROA $\Delta$ -value

Comparing the ROA spectra of disaccharide D-maltose (Fig. 6.15) with its monomer D-glucose (Fig. 3.7) reveals that the two ROA spectra are very similar in the fingerprint and above  $1200\text{cm}^{-1}$  regions. Both give a couplet at  $\sim 1325\text{cm}^{-1}$  which is associated with the  $\text{CH}_2\text{OH}$  group and the ROA feature between  $960\text{cm}^{-1}$  and  $1200\text{cm}^{-1}$  is almost identical.

The most distinct ROA feature between the two molecules is a couplet centered at  $\sim 910\text{cm}^{-1}$  positive at lower frequency and negative at higher side which appears in maltose. No such couplet appears in the ROA spectrum of D-glucose in this region. So this couplet for maltose may be associated with the glycosidic linkage. Raman bands for  $\alpha$ -D-glucose at  $\sim 845\text{cm}^{-1}$  and  $\sim 914\text{cm}^{-1}$  have been assigned to vibrational modes that involves contribution from C-1'-H deformation.<sup>15</sup> The C-O-C stretch that involve the  $\alpha$ -(1-4) linkage in maltose appears in the  $920\text{-}960\text{cm}^{-1}$  region and this couplet is attributed to the interaction of the C-1'-H deformation

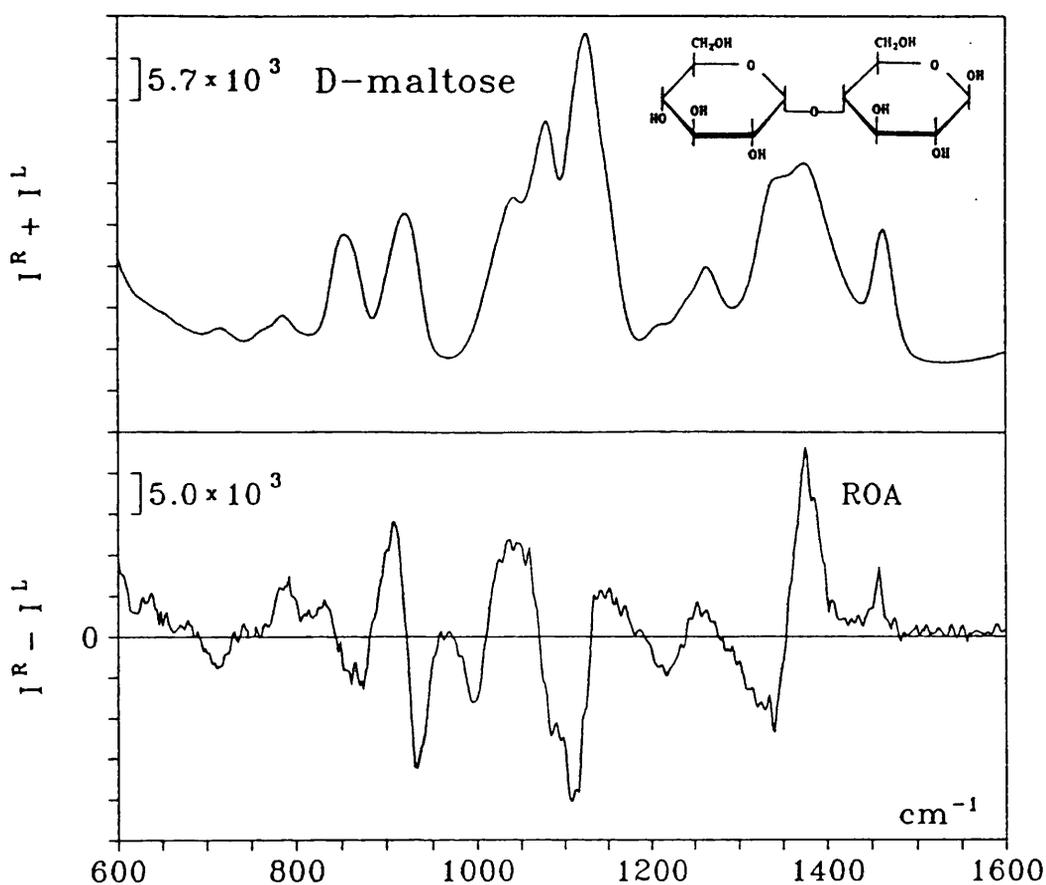


Fig. 6.15 Raman and ROA spectra of D-maltose in water.

and the glycosidic C-O-C stretch, probably with some contribution from the C-4-H deformation.

The ROA couplet at  $\sim 910 \text{ cm}^{-1}$  also appears in the spectrum of maltotriose (not shown here) which supports its assignment to the glycosidic linkage and indicates that the two glycosidic linkages in maltotriose are similar to each other and to that in maltose.

The ROA couplet becomes enormous with a  $\Delta$  value of  $10^{-2}$  in  $\alpha$ -D-cyclodextrin which is composed of six glucosyl units joined by six

glycosidic bonds ( Fig. 3.10 ). The  $\Delta$  value is one order of magnitude larger than the largest usually encountered. The enormous  $\Delta$  value of this couplet in cyclodextrins indicates that they originate in coupled vibrational modes of the glycosidic links delocalized around the cyclodextrin ring. The potentially most valuable result is the identification of clear ROA features associated with the glycosidic linkage, which appear to be a good probe of the conformation of the polysaccharides.

Estimation of the ROA  $\Delta$  value of these carbohydrates reveals that it could be one or two orders of magnitudes difference among stereoisomers of monosaccharides. The ROA  $\Delta$  value seems to be a manifestation of the sugar ring structure. For example, as mentioned in section 6.3.4, the ROA  $\Delta$ -value of D-lyxose and D-arabinose are much larger than the similar pentopyranose D-xylose due to the existence of a considerable proportion of the  $1C_4$  conformation of the pyranoid ring (remembering that all the substituents on the pyranoid ring of D-lyxose are in axial orientation in the  $1C_4$  conformation which makes the substituents 'crowded' whereas they are virtually all in the equatorial position in D-xylose). The spectral quality of methyl-D-pyranosides are much better than that of corresponding pyranoses, which in turn reflect the larger ROA  $\Delta$ -values of the pyranosides. Table 6-2 summarizes the largest ROA  $\Delta$ -values of typical saccharides. It can be seen from the Table 6-2 that the ROA  $\Delta$ -values correlate well with the sugar ring conformation. The more rigid the ring structure, the stronger the ROA. The largest  $\Delta$ -value result from the glycosidic linkage in  $\alpha$ -D-cyclodextrin and the least is in D-ribose and D-talose which exist as multiple equilibria between the ring tautomerism

Table 6-2. Correlation between the largest ROA  $\Delta$  value and saccharide ring structure.

saccharides	sugar ring conformation	ROA $\Delta$ value
$\alpha$ -D-cyclodextrin	pyranose C1	$\sim 1 \times 10^{-2}$
1,6-anhydro- $\beta$ -D-glucopyranose	pyranose C1	$\sim 5 \times 10^{-3}$
D-lyxose	pyranose 1C $\longleftrightarrow$ C1	$\sim 1 \times 10^{-3}$
D-arabinose	pyranose 1C $\longleftrightarrow$ C1	$\sim 1 \times 10^{-3}$
methyl- $\alpha$ -D-pyranosides	pyranose C1	$\sim 1 \times 10^{-3}$
D-glucose	pyranose C1, ( $\beta$ anomer 64%)	$\sim 5 \times 10^{-4}$
D-ribose	pyranose $\longleftrightarrow$ furanose	$\sim 1 \times 10^{-4}$
D-talose	pyranose $\longleftrightarrow$ furanose	$\sim 1 \times 10^{-4}$

and anomerisation, so that the structure of the sugar ring is much more complex and exists in a mixture of  $\alpha$ -,  $\beta$ -D-pyranose and  $\alpha$ -,  $\beta$ -D-furanose.

#### 6.4. Conclusions

The carbohydrate ROA spectra reported and discussed here demonstrate that ROA is capable of providing detailed stereochemical information on carbohydrates in aqueous solution inaccessible to other spectroscopic methods. The ample information content of carbohydrate ROA spectra contrasts dramatically with the paucity of information available from conventional ECD studies. If the many correlations between ROA spectra and carbohydrate structure can be established, they would be invaluable to carbohydrates stereochemistry. On the other hand, ECD can

study much more dilute solutions than ROA; but this limitation will soon be relieved as further anticipated developments in ROA instrumentation will provide greater sensitivity.

The delocalization of many of the normal modes over the many C-C and C-O linkages, which leads to difficulties in the interpretation of conventional Raman spectra of carbohydrates and has prevented its widespread application in carbohydrate chemistry,<sup>25</sup> is in fact a pre-requisite for large vibrational optical activity and leads, as we have seen, to characteristic ROA features which are immediately discernible. Furthermore, the presence of C-O-C linkages has been shown previously to be particularly favourable for large ROA intensities.<sup>26-28</sup>

The information content of an ROA band is a function of its frequency, sign, intensity and shape. So far, only the ROA band frequency and sign have been used for correlations with stereochemical features: the usefulness of the dimensionless  $\Delta$ -value as a measure of ROA intensity has not been explored much. However, as mentioned above, the ROA  $\Delta$ -value appears to be a valuable source of information about the ring structure and conformation of monosaccharides with the  $\Delta$ -value of similar monosaccharides differing by up to one order of magnitude (also the  $\Delta$ -value of ROA bands associated with the glycosidic linkage in cyclodextrins is two orders of magnitude larger than the other common ROA features ). However, the complexity of many of the parent carbohydrate Raman bands require that band decomposition procedures will have to be employed and the spectral quality reproducibility needs to be further improved in order to obtain precise estimates of the ROA band

$\Delta$ -values.

Just as *ab initio* ROA computations on the archetypal amino acid alanine have provided a valuable springboard for understanding the origin of important ROA features in peptides and proteins,<sup>10</sup> so an *ab initio* ROA computation on a simple monosaccharide such as arabinose or glucose will help develop an understanding the origin of ROA spectra of carbohydrates at a fundamental level rather than simply relying on empirical correlations. Anticipated developments in the computation of the required optical activity tensor derivatives should soon render small monosaccharides accessible to such *ab initio* ROA calculation.

## REFERENCES

1. R. J. Sturgeon, *Chemistry in Britain*, **26**, 663, 1990.
2. S. J. Angayal, *Angew. Chem. International Edition English*, **8**, 157, 1969.
3. D. Parker, *Chem. Rev.*, **91**, 1441, 1991.
4. S. F. Mason, *Molecular Optical Activity and Chiral Discrimination*, Cambridge University Press, Cambridge, 1982.
5. W. C. Johnson, *Adv. Carbohydr. Chem. and Biochem.*, **45**, 73, 1987.
6. M. Mathlouti and J. L. Koenig, *Adv. Carbohydr. Chem. and Biochem.*, **44**, 7, 1986.
7. a, M. G. Peterlini, T. B. Freedman and L. A. Nafie, *J. Am. Chem. Soc.*, **108**, 1389, 1986. b, C. T. Tummalapalli, D. M. Back, P. L. Polavarapu, *J. Chem. Soc. Faraday. Trans 1.*, **84**, 2585, 1988.
8. L. D. Barron, A. R. Gargaro and Z. Q. Wen, *Carbohydr. Res.*, **210**, 39, 1990.
9. P. L. Polavarapu, *J. Phys. Chem.*, **94**, 8106, 1990.
10. L. D. Barron, A. R. Gargaro, L. Hecht and P. L. Polavarapu, *Spectrochimi. Acta*, **47A**, 1001, 1991. *ibid*, **48A**, 261, 1992.
11. L. D. Barron, A. R. Gargaro, Hecht, P. L. Polavarapu and H. Sugeta, *Spectrochimi. Acta*, **48A**, 1051, 1992.
12. L. D. Barron, A. D. Buckingham, *J. Am. Chem. Soc.*, **96**, 4769, 1974.
13. L. D. Barron, *Molecular Light Scattering and Optical Activity*, Cambridge University Press, Cambridge, 1982.
14. S. F. Barker, E. J. Bourne, M. Stacey and D. H. Whiffen, *J. Chem. Soc.*, 174, 1954.
15. J. J. Cael, J. L. Koenig and J. Blackwell, *Carbohydr. Res.*, **32**, 79, 1974.

16. P. D. Vasko, J. Blackwell and J. L. Koenig, *Carbohydr. Res.*, **23**, 407, 1972.
17. P. D. Vasko, J. Blackwell and J. L. Koenig, *Carbohydr. Res.*, **19**, 297, 1971.
18. A. F. Bell, Z. Q. Wen, L. D. Barron and L. Hecht, to be published.
19. R. G. Nelson, W. C. Johnson Jr, *J. Am. Chem. Soc.*, **94**, 3342, 1972.
20. A. D. Bruyn and M. Anteunis, *Carbohydr. Res.*, **19**, 297, 1971.
21. V. V. Sivchik, R. G. Zhabankov, *Zh. Prikl. Spektrosk.*, **32**, 1056, 1980.
22. N. K. de Vries and H. M. Buck, *Carbohydr. Res.*, **165**, 1, 1987.
23. M. Mathlouti, A. M. Seuvre, J. L. Koenig, *Carbohydr. Res.*, **122**, 31, 1984.
24. P. Carmona, M. Molina, *J. Raman Spectrosc.*, **21**, 385, 1990.
25. P. R. Carey, *Biochemical Application of Raman and Resonance Raman Spectroscopies*, Academic Press: New York, 1982.
26. L. D. Barron, P. L. Polavarapu, *Mol. Phys.*, **65**, 659, 1988.
27. T. M. Black, P. K. Bose, P. L. Polavarapu, L. D. Barron and L. Hecht. *J. Am. Chem. Soc.*, **112**, 1479, 1990.
28. P. K. Bose, P. L. Polavarapu, L. D. Barron and L. Hecht, *J. Phys. Chem.*, **94**, 1734, 1990.

# Publications

## A. Publications in refereed journals

### 1. Vibrational Raman optical activity of peptides and proteins,

L. D. Barron, A. R. Gargaro and **Z. Q. Wen**, *J. Chem. Soc., Chem. Commun.*, 1034, (1990).

### 2. Vibrational Raman optical activity of cyclodextrins,

L. D. Barron, A. R. Gargaro, **Z. Q. Wen**, D. D. MacNicol and C. Butters, *Tetrahedron: Asymmetry*, 1, 513, 1990.

### 3. Vibrational Raman optical activity of carbohydrates,

L. D. Barron, A. R. Gargaro and **Z. Q. Wen**, *Carbohydr. Res.*, 210, 39, 1991.

### 4. Vibrational Raman optical activity of proteins,

L. D. Barron, **Z. Q. Wen** and L. Hecht, *J. Am. Chem. Soc.*, 114, 784, 1992.

### 5. Raman optical activity instrument for biochemical studies,

L. Hecht, L. D. Barron, A. R. Gargaro, **Z. Q. Wen** and W. Hug, *J. Raman Spectrosc.*, 23, 401, 1992.

## B. Publications in conference proceedings

### 1. Raman optical activity of biological molecules,

L. D. Barron, A. R. Gargaro, L. Hecht, **Z. Q. Wen** and W. Hug, in 'Laser Application in Life Sciences' edited by S. A. Akhmanov and M. Y. Poroshina, Moscow, SPIE 1403, 66, 1990.

**2. Raman optical activity of biological molecules,**

L. D. Barron, L. Hecht, A. R. Gargaro, **Z. Q. Wen** and W. Hug, in *"Spectroscopy of Biological Molecules"*, edited by R. E. Hester and R. B. Girling, The Royal Society of Chemistry, Cambridge, 117, 1991.

**3. Vibrational Raman optical activity of carbohydrates,**

**Z. Q. Wen**, L. D. Barron and L. Hecht, in *6th European Symposium on Carbohydrate Chemistry*, The Royal Society of Chemistry, Edinburgh, A7, 1991.

**4. Vibrational Raman optical activity of biological molecules,**

L. D. Barron, L. Hecht, **Z. Q. Wen**, S. J. Ford and A. F. Bell, *13th International Conference on Raman Spectroscopy*, edited by W. Kiefer, John Wiley and Sons, Wurzburg, Sep, 1992.

**5. Raman optical activity of nucleosides and nucleotides,**

L. Hecht, L. D. Barron, **Z. Q. Wen** and S. J. Ford, *13th International Conference on Raman Spectroscopy*, edited by W. Kiefer, John Wiley and Sons, Wurzburg, Sep. 1992.

**6. Raman optical activity of biological molecules,**

L. D. Barron, L. Hecht, **Z. Q. Wen**, S. J. Ford and A. F. Bell, *4th International Conference on 'Laser Application in Life Sciences'*, edited by J. Korppi-Tommola, Jyvaskyla, Finland, Sep. 1992.

**7. Raman optical activity of enzymes,**

S. J. Ford, L. D. Barron, L. Hecht, A. Cooper and **Z. Q. Wen**, *4th International Conference on 'Laser Application in Life Sciences'*, edited by J. Korppi-Tommola, Jyvaskyla, Finland, Sep. 1992.

**C. To be published**

**1. Raman optical activity of monosaccharides,**

Z. Q. Wen, L. D. Barron and L. Hecht, *J. Am. Chem. Soc.*, in press.

**2. Raman optical activity of enzymes,**

L. D. Barron, A. Coopers, S. J. Ford, L. Hecht and Z. Q. Wen, *Faraday Discussion, Chem. Soc. No# 93*, The Royal Society of Chemistry, Cambridge, 1992, in press.

**3. Raman optical activity of L-alanine oligomers,**

Z. Q. Wen, L. D. Barron, L. Hecht and S. J. Ford, *Biopolymers*, in preparation.