NOVEL RAMAN STUDIES OF THE CONFORMATION AND DYNAMICS OF BIOLOGICAL MOLECULES

By Steven John Ford

A Thesis presented in partial fulfilment for
the degree of Doctor of Philosophy in the
Faculty of Science of the
University of Glasgow

Chemistry Department

May 1995

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

To Sandra

ACKNOWLEDGEMENTS

I would like to express my deepest thanks to Prof. Laurence D. Barron and Dr. Alan Cooper, for acting as my supervisors. Their relaxed, yet very professional, attitude toward research has been a great encouragement to me over the last 3 years.

I would like to thank the friends whom I have worked beside during my Ph.D.: Stephen Robertson, Alan M^{c}Alpine, Michelle Lovat, Deborah M^{c}Phail, Margaret Nutley, Gary Wilson, Lutz Hecht and Alasdair Bell.

I would like to thank my extended family, particularly my parents and my wife, and my friends, John Candy, Duncan Brown, Ian Futter and Andrew Munro for all the support they have given me over the last few years.

ABBREVIATIONS

BabBAL	Baboon α-lactalbumin
BAL	Bovine α-lactalbumin
BSA	Bovine serum albumin
CCD	Charged coupled device
CID	Circular intensity difference
DCP	Dual circularly polarized
DTT	Dithiolthreitol
ICP	Incident circularly polarized
EOM	Electro-optic modulator
FTIR	Fourier transform infra-red
HEWL	Hen egg white lysozyme
ipb	in-plane bend
IR	Infra-red
MD	Molecular dynamic
NAG	N-acetyl glucosamine
NOE	Nuclear Overhauser effect
NMR	Nuclear magnetic resonance
opb	out-of-plane bend
ORD	Optical rotatory dispersion
PLL	Poly-L-lysine
PrOH	Propanol
ROA	Raman optical activity
TFE	Trifluoroethanol
UVCD	Ultra-violet circular dichroism
VCD	Vibrational circular dichroism
VOA	Vibrational optical activity

CONTENTS

DIAGRAMS

TABLES

ABSTRACT

Raman optical activity (ROA) is developing into a new, and unique, form of biological spectroscopy and this thesis endeavours to determine the new insights that ROA provides into the conformational and dynamic properties of peptides and proteins.

The peptide ROA analysis used small oligopeptides to obtain a clear idea of the fundamental aspects of polypeptide ROA. The ROA signal originating from the vibrations of the amide link were scrutinized particularly closely using the alanyl peptide series Ala_2 to Ala_5. The amide III vibration (C_α-H and N-H deformations) proved to be the easiest, and most reliable, band to measure experimentally. The coupling of the C_α-H and N-H deformations can be seen clearly in both the Raman and the ROA spectra. The amide III Raman and ROA analysis concurs with recent normal mode calculations. As the alanyl series expands delocalization of the amide III band can be seen. It is clear that the delocalization is dependent upon the N-H deformations because no changes are seen in the C_α-H deformations as the deuterated alanyl peptide (with low frequency N-D deformations) series expands.

Other peptides were also studied by ROA, and this gave a greater insight into the peptide ROA observed from all the molecules under examination. Several alanyl ROA assignments were confirmed by the work. The amide III ROA differed between peptides, indicating conformational changes between the molecules. However, the C_α-H deformations from deuterated peptides were seen to give similar ROA band envelopes in all the peptides studied, indicating again the conformational sensitivity of the amide III ROA signal and the importance of the N-H deformation. Significant amide I and amide II Raman and ROA characteristics were also observed.

The protein studies concentrated on hen egg white lysozyme (HEWL) and bovine α-lactalbumin (BAL), two well studied and understood proteins. The Raman and ROA spectra of the native, denatured, inhibitor bound, deuterated and partially deuterated states of HEWL have been observed, as have the native, molten globule and metal free states of BAL. It has become apparent that the amide I vibration produces an ROA couplet in most protein states, while the amide III ROA is very sensitive to protein conformation. Several of the ROA bands associated with the amide III vibration have been assigned, and particularly interesting are the bands that are sensitive to rigid, dynamic or disordered structural features. A simplistic theoretical model, that explains some of the ab initio ROA calculations and the general 'couplet' nature of the protein amide III ROA signal, is outlined in the final chapter.

Although the ROA signals from proteins cannot be assigned with complete confidence, it is clear that ROA is rapidly developing into a new form of biological spectroscopy, with the exciting possibility of probing the dynamic aspects of protein structure.

Chapter 1

Introduction

1.1 BACKGROUND

1.1.1 Objective

Proteins are the building blocks and workhorses of life. Yet the understanding of their function still remains one of the greatest challenges of biological chemistry. Raman optical activity is in a key position to make a contribution to this challenge. It is a unique form of spectroscopy: its signals depend upon molecular properties that no other technique can analyse. Therefore, it is essential to determine what insight into biomolecular function protein Raman optical activity can give. This is the objective of this thesis.

1.1.2 Introduction

Protein structure and function are intimately connected. In the last 30 years many protein structures have been deduced by X-ray crystallography and NMR. (The latter being used to determine protein solution structure.) However, it is becoming increasing apparent that the dynamic behaviour of biomolecules in solution also plays a major role in their function (Cooper, 1980; Brooks et al., 1988; McCammon and Harvey, 1987). Being able to 'see' a three dimensional structure is no longer good enough, protein function depends on a 'four dimensional' structure: how protein conformation changes with time; how protein molecules move and act. (Tailor's dummies rarely move, act and behave like humans!)

X-ray diffraction and NMR begin to show their limitations when this extra dimension of time is considered: X-ray diffraction requires a virtually static protein structure which is 'locked' into a crystal and NMR structure determination becomes very difficult when the protein undergoes large dynamic motion. Therefore, a broad range of physical and chemical methods have been devised examine protein behaviour and it is in this category that vibrational spectroscopy belongs.

Vibrational spectroscopy has been used as a biomolecular tool for quite some time, but, with protein crystallography producing tangible 3D structures, vibrational spectroscopy, and other 'non-pictorial' techniques, were sidelined. These techniques are now re-emerging as their advantages are being reassessed, and with them is a new breed of vibrational spectroscopy - vibrational optical activity (VOA).

1.2 VIBRATIONAL OPTICAL ACTIVITY

1.2.1 Introduction

Vibrational optical activity (VOA) is a chiroptical extention of vibrational spectroscopy in which ordinary Raman or IR experiments are enhanced by the use of circularly polarized light. The consequence of this combination is that the chiroptical vibrational spectrum is significantly more sensitive to molecular stereochemistry. Moreover, this sensitivity is extended beyond traditional chiroptical techniques (such as ultraviolet circular dichroism) because the VOA signal depends on the encompassing of a chiral structure by a molecular vibration; any molecule has 3N-6 normal modes of vibration but will only have an electronic transition if the appropriate chromophore is present. So, on one hand we have the advantage of a vibrational spectroscopy - vibration from all parts of all molecules - and on the other we have the advantage of a chiroptical

vibrational spectroscopy - great sensitivity to molecular stereochemistry with a localization of the signal to chiral structures.

1.2.2 Experimental VOA techniques

As a conventional vibrational spectrum can be measured by Raman or IR, both of which depend on different molecular properties, so a VOA spectrum can also by determined by two different methods:

1) Raman optical activity (ROA) - the difference in the Raman scattering with circularly left and right polarized incident and/or scattered light.

2) Vibrational circular dichroism (VCD) - the difference in absorption of circularly left and right polarized IR radiation.

These two forms of VOA are very different and they both depend upon very different molecular properties; therefore ROA and VCD are considered to be independent, and complementry, techniques. Circular dichroism is based upon the polarization characteristics of light *transmitted* through an optically active medium and involves the interference of unscattered and forward scattered light waves. This is shown diagramatically in Figure 1.1(a), which illustrates a simple two group model (which is made up of two identical chemical groups held in a twisted chiral arrangement). In order to determine the chirality of the molecule via circular dichroism a photon (or the electromagnetic field associated with that photon) must pass from one group to another before being scattered in the forward direction. Rayleigh and Raman optical activity are based upon the polarization characteristics of *scattered* light. They are still interference phenomena, but the interference occurs between two light waves scattered independently from the two groups (as shown in Figure 1.1(b)).

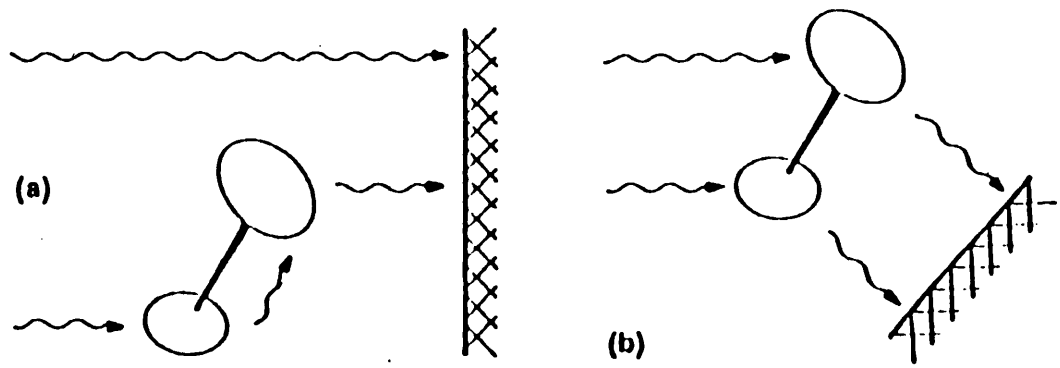


Figure 1.1: The photon scattering picture of the generation of (a) circular dichroism and (b) Rayleigh and Raman optical activity by a simple chiral two-group model. (Taken from Barron and Hecht, 1994).

1.2.3 VOA as a biochemical tool.

The VOA signal produced by a vibrational mode depends upon the incorporation of a chiral structure within that vibrational mode. Therefore, a VOA signal will only be observed for 'optically active' vibrations. So when we turn to protein ROA our attention is naturally drawn towards the chiral units within a protein structure - amino acids.

There are 20 naturally occuring amino acids and all of them, with the exception of glycine, contain a chiral C_α. However, two of the amino acids, threonine and isoleucine, have a second chiral centre, on the C_β.

When amino acids are polymerized into a protein structure the C_α's are incorporated into the backbone chain, leaving very few chiral centres on the sidechains. (The sequence of hen egg white lysozyme shows that 90% of the chiral centres are located in the protein backbone.) This natural focussing of chiral centres on a particular part of the structure combined with the localization of VOA signals onto chiral centres results in a technique that is very sensitive to the molecular conformation of the protein backbone. Chirality can also be generated by chemical groups held

in a twisted conformation within the protein structure, such twisting is believed to contribute to protein VCD (Malon, 1994).

The nature of normal mode vibrations means that several vibrations will encompass a single chiral centre; consequently VOA techniques will have several bands which contain the same structural information. ROA puts this advantage to good use because it has a far broader spectral range than VCD. Ultraviolet circular dichrosim (UVCD), which is the traditional chiroptical biochemistry tool, depends upon the two electronic transitions of the peptide link.

Many other biomolecular spectroscopies, for example some NMR studies, near-UVCD and fluorescence, do not actually determine backbone conformation. Instead they rely on the aromatic residues within the protein structure to act as intrinsic probes and infer backbone conformational changes from the changes in the aromatic environment. VOA uses a different, and more direct, approach, sampling the backbone conformation directly.

Conventional Raman and IR rely solely on the frequency and intensity of vibrational modes to deduce protein secondary structure. However, it is becoming increasingly apparent that this approach is not suitable because the 'conformationally sensitive' vibrational modes are also affected by other factors, such as sidechain structure and geometric irregularity (Krimm and Bandekar, 1986; Krimm and Reisdorf, 1994). VOA provide another parameter, which will hopefully be less sensitive to these other factors (Keiderling, 1994).

VCD has been successfully used on peptides and proteins for over a decade, allowing several different aspects of experimental procedures (and theoretical understanding) to be developed. Experimentally, the amount of

protein required for a successful VCD result is much smaller than for ROA because of the smaller sample volume and lower concentration in the former. Many of the intrumental problems involved with biological VCD have been solved and so acquiring a reliable spectrum is relatively straightforward.

VCD compares favourably to its high frequency cousin, UVCD. The characteristic frequencies of particular secondary structural elements in a UVCD spectrum overlap, whereas with VCD they are resolved. VCD is also more sensitive to shorter secondary structural elements than UVCD because the coupling between the vibrational transitional moments occurs over shorter regions than the coupling for the electronic transitional moments (Birke et al., 1992). One the other hand attempts at using a statistical factor analysis on protein VCD results have shown that VCD is no more reliable than UVCD for secondary structure prediction (Pancoska et al., 1990, 1994). The statistical analysis is now being extended to try and predict supersecondary structure (Pancoska et al., 1994).

One of the advantages of Raman spectroscopy over IR spectroscopy is that the Raman intensity of the H_2O bending vibration is very small: whereas Raman spectroscopy can determine bands that coincide with the 1660 cm^{-1} H_2O bend, IR has to rely on solvent subtraction techniques. The same situation applies to D_2O except that the vibration is now at 1200 cm^{-1}. Therefore, ROA spectroscopy is equally applicable to proteins solvated in H_2O or D_2O and this is important when studying amide hydrogen-deuterium exchange. VCD does not have this advantage and relies on deuterated proteins. However, because complete deuteration requires denaturation and regaining the native conformation is not always straightforward, VCD measurements are done on partially deuterated proteins. Recently VCD spectra have been obtained from proteins in H_2O (Keiderling et al., 1994).

1.2.4 ROA strategies.

ROA is a blanket title which covers several experimental strategies. Unlike VCD, the nature of the ROA measurement can be varied quite widely. Different ROA stategies produce different ROA spectra because when the ROA strategy is changed different molecular properties contribute to the observed ROA signal. One experimental parameter that can be altered is scattering angle and, although any scattering angle between 0^{o} and 180^{o} could be employed, the measurements are usually restricted to forward scattering (0^{o}), right angle scattering (90^{o}) and backscattering (180^{o}). For reasons which are further alluded to in the theoretical section, the protein ROA experiments done in this thesis are based on a backscattering experimental approach.

Until recently the only ROA strategy that was used was incident circularly polarized (ICP) ROA. This form of ROA, as the name implies, was based on alternating the incident laser light between circularly left and right polarized forms. The ICP ROA experiment is the strategy that has been adopted at Glasgow and is used for all the measurements in this thesis. The technique can perhaps be visualised by using a cricketing analogy - if a bowler alternates the spin of the ball he throws he will be able to change the behaviour of the ball once it has been hit: similarly a right handed batsman, who represents the chiral sample here, will have a different effect on the deflection of the spinning ball(s) than a left handed teammate.

The new ROA development is the adoption of polarized modulation in the beam of scattered light - scattered circularly polarized (SCP) ROA (Hecht et al., 1991). The ICP and SCP can be combined to produce a dual circularly polarized (DCP) ROA experiment which, naturally, can have two forms depending upon the combinations of circularly polarized light used.

The first form, in-phase DCP (designated DCP_I), is where the polarized forms of the incident and scattered beams are both the same. Out-of-phase DCP (designated DCP_{II}) is where the polarized forms of the incident and scattered beams are different; that is right circularly polarized incident light is used when the scattered light analyser is set to left circularly polarized light, and vice versa. All four strategies are shown in Figure 1.2.

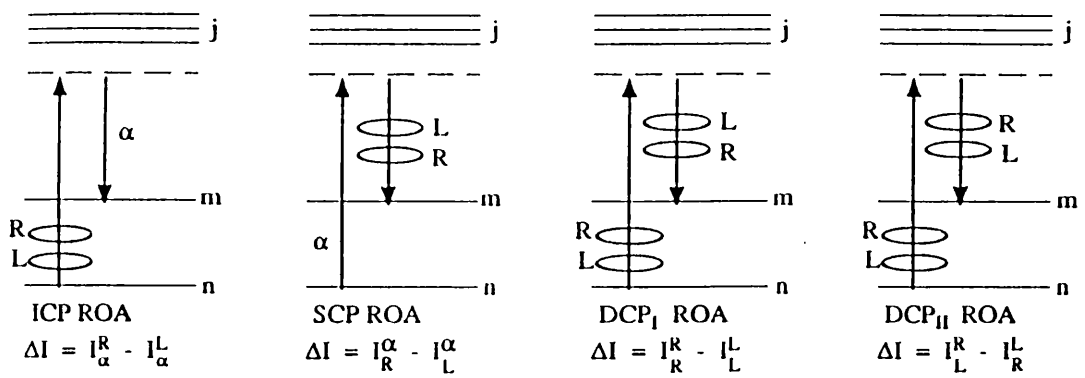


Figure 1.2: Different ROA strategies. (Taken from Nafie, 1991)

1.2.5 Vibrational spectroscopy of proteins.

This final section discusses the classification of the vibrational signals that occur in protein spectra and how conventional vibrational spectroscopy uses these vibrations.

For the purpose of vibrational spectroscopy protein structure can be divided into two main components, sidechain residues and amide links, each of which will contribute to the observed spectrum. However, a protein, being a macromolecule, will have a very large number of normal vibrational modes and consequently, it is impossible to discern particular vibrations, unless the signal from those vibrations is intense enough to stand above, and be distinct from, the rest. Furthermore, because proteins have so many

vibrations at a particular frequency, the observed band will be a summation of several vibrations from various components within the protein.

The sidechain residues that appear in a Raman spectrum are usually the aromatic residues, tryptophan, phenylalanine and tyrosine and the sulphur containing residues, methionine and cysteine. Some of the vibrations of these residues are sensitive to the environment that the residue is in and can act as intrinsic conformational probes within the protein.

The second component of protein structure, the amide links, have their own set of vibrations. The amide vibrations have been extensively studied to determine the their dependence on the backbone conformation. The amide vibrations are outlined in Table 1.1.

Table 1.1: The vibrational modes of the amide links. (s; stretch: ip; in plane bend: b; bend: t; torsion: op; out of plane bend.)

Amide mode	Frequency (cm^{-1})	Vibrational components
Amide I	1640-1680	C=Os with C-Ns and N-Hip
Amide II	~1560	N-Hip and C-Ns
Amide III	1230-1300	N-Hip and C-Hb
Amide IV	various	C=Oip
Amide V	~725	C-Nt and N-Hop
Amide VI	various	C=Oop
Skeletal stretch	880-960	Backbone C-Cs and C-Ns

The Raman and IR bands from the amide vibrations depend on the structure of the protein being examined. Several methods have been proposed to qualitatively determine the secondary structure contents of proteins from the amide vibrations with varying degrees of success. (A recent review of these methods has been published by Bandekar (1992).) However, these techniques have probably not been widely adopted because

of the experimental difficulties in obtaining a good Raman or IR spectrum as opposed to a good UVCD spectrum. This state of affairs may changes as solvent subtraction, deconvolution, Fourier transform and infra red Raman techniques improve.

Chapter 2

ROA methods and theory

2.1. ROA EXPERIMENTAL METHODS.

2.1.1 Introduction

This section provides a brief description of the ROA spectrometer used for the measurements in this thesis. A full description can be found in a recent paper (Hecht et al., 1992).

The design of the spectrometer is based on several year's study into the experimental and theoretical aspects of ROA collection strategies, and also incorporates some of the latest optical technology (single stage spectrograph, CCD dectector, holographic notch filter etc.). The collection strategy that was used is that of backscattering ICP ROA. Although not widely used for Raman spectroscopy, the ROA signal is strongest in the backscattering direction.

2.1.2 The ROA spectrometer.

The optical layout of the spectrometer is shown in Figure 2.1 and each individual component described systematically below.

The laser is a continuous-wave Spectra-Physics argon ion laser (Stabilite 2016), set to 514.5 nm. The first aperture was used to suppress plasma background radiation from the laser. The polarizer enhanced the linear polarization of the laser light.

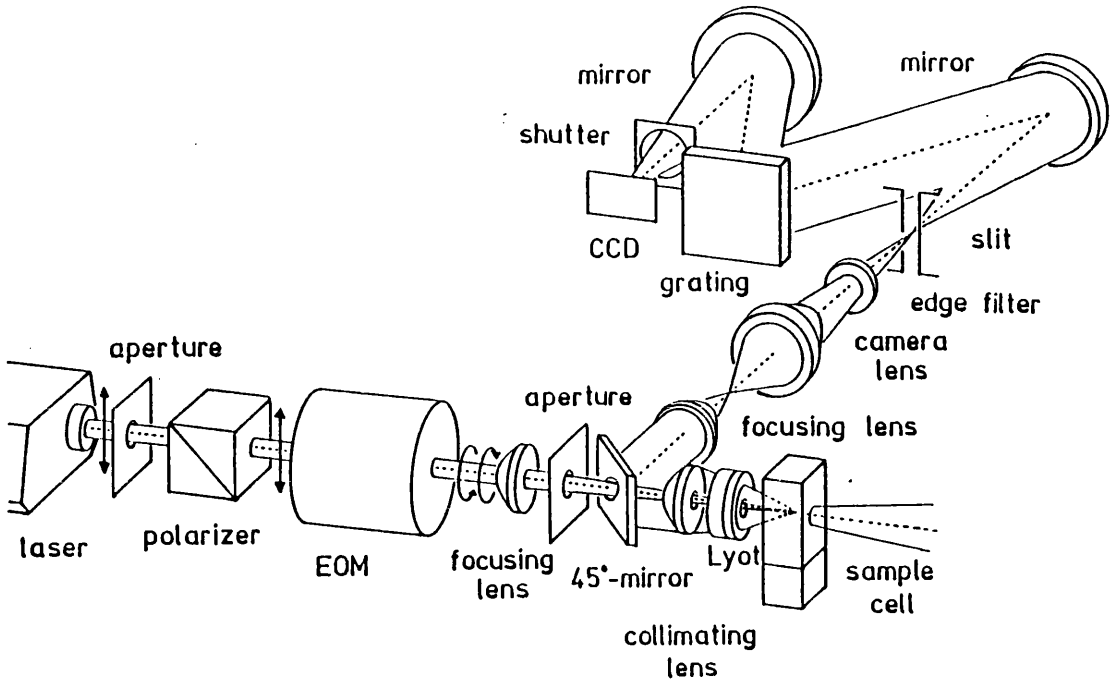


Figure 2.1: The optical layout of the backscattering ICP ROA spectrometer. (Adapted from Hecht et al., 1992)

The electro-optic modulator (EOM) is a Pockels cell (Electro-Optic Developments Model 850-16), which consists of a crystal of potassium dideuterium phosphate sandwiched between two gold ring electrodes. When a high voltage (~3kV) is applied along the direction of the laser beam the crystal acts as a quarter-wave retarder, converting linearly polarized light to circularly polarized light: reversing the sense of the high voltage switches the circular polarization of the transmitted light from right to left or vice versa. The EOM is maintained at a constant temperature, by a Electro-Optical Developments Model TC 15 temperature controller, because the birefringence properties of the Pockels cell are temperature dependent. The voltage required for the optical modulation is provided by a high voltage linear differential amplifier (LEYSOP, Model LDA 50 Series 5000). The resulting circularly polarized light is focussed onto the sample cell and passes through a 1 mm diameter aperture.

The position of the laser focussing lens can be altered and this was occasionally used to defocus the laser beam at the sample. This reduced the light flux at the sample, so avoiding photolytic degradation, but still maintained the overall intensity of light passing through the sample.

The next three components were housed in a block placed just behind the sample cell and are shown in Figure 2.2. All three components have centrally placed holes (2mm diameter) to let the focused circularly polarized light shine directly onto the sample cell.

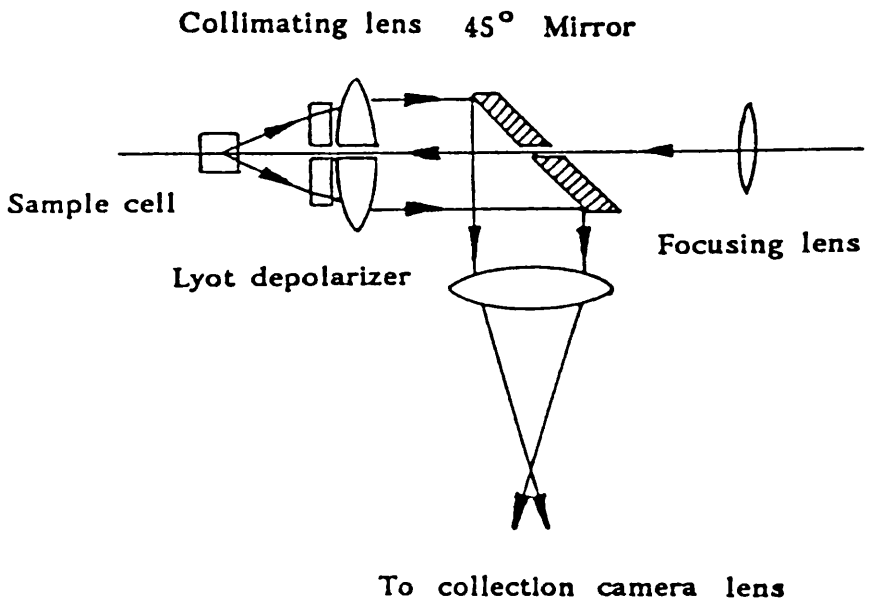


Figure 2.2: The backscattering block.

The cone of backscattered light that falls on the Lyot depolarizer, is collimated by the lens and reflected into the collection optics by the 45° angle mirror. The Lyot depolarizer is essential because without it the mirror induces large ROA artifacts in the spectrum. (A mirror does not reflect different forms of polarized light equally, so the Lyot depolarizes the scattered light before it is reflected.)

The sample cell is a specially manufactured quartz microfluorescence cell (Optiglass) with a pathlength of 5mm and a maximum sample volume of 300 μl (although the usual sample volume for proteins was 100-150 μl).

The collection optics consist of a focusing lens and a camera lens which focuses the collected depolarised light, through a holographic notch filter, onto the spectrograph slit (for protein ROA the latter was usually set at 120 microns). The holographic notch filter removes the Rayleigh scattered light, allowing only the inelastically (Raman) scattered light to enter the spectrograph. The spectrograph is a single stage Jobin Yvon HR-250S spectrograph, with an ion-etched 1200 grooves mm^{-1} grating providing the diffraction.

The detection system is based on a backthinned charged coupled device (CCD) manufactured by Wright Instruments. The backthinned CCD consists of 385 columns x 578 rows and is oriented such that the short axis is parallel to the spectrograph slit. The CCD exposure is controlled by an internal shutter within the CCD housing and is maintained at a temperature of 200K by a Peltier cooler reduces the dark counts to a negligible value. After exposure the CCD is vertically binned so that, although it is a 2D array of detection cells, it acts as a linear array of long narrow vertical pixels.

The whole experiment is computer controlled, with the only manual input once the ROA acquistion has started being to 'fine tune' the high voltage output of the amplifier (this allows greater artifact control during the acquisition). The CCD operation and high voltage switching is regulated by a Wright Instruments AT computer interface and a Dell 325D AT PC using the Lab Calc spectral acquistion and manipulation software (Galactic Industries). The relationships between the various hardware components are shown in Figure 2.3.

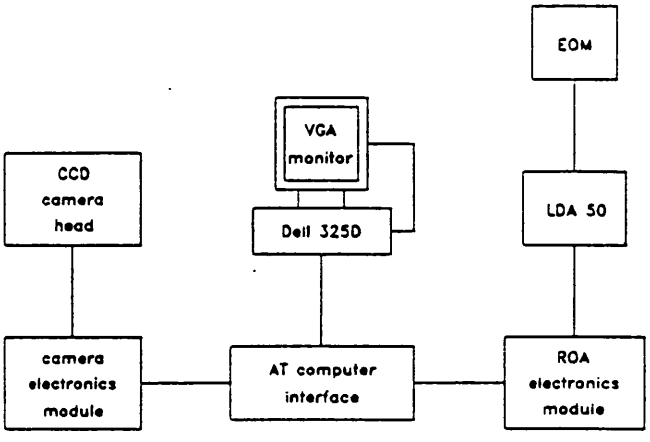


Figure 2.3: Diagram of the hardware connections between the various electronic components of the ROA spectrometer. (Adapted from Hecht et al., 1992)

The operation of the ROA spectrometer, both manual and electronic, are shown in the flow diagram in Figure 2.4.

2.1.3 Experimental parameters.

The laser power was usually set to 1.25 - 1.4 W. The Raman and ROA frequencies are accurate to ±2 cm^{-1}. The frequency range is about 1000 cm^{-1} so in order to acquire a spectrum 400 to 1800 cm^{-1} the ROA spectrum has to be recorded in two separate sections.

Once the ROA spectrum has been obtained it is some times cosmetically altered. This is done for several reasons: the CCD is sensitive to cosmic rays and when one strikes the CCD it produces a very large 'spike' in the ROA spectrum; in some ROA spectra artifacts can appear on one band leaving an otherwise good ROA spectrum dominated perhaps by one large S-artifact; the acquisition of protein ROA can take between 12 and 24 hours and in that time baseline drifts occur, usually at the high frequency end of the spectrum as the fluorescence background decreases.

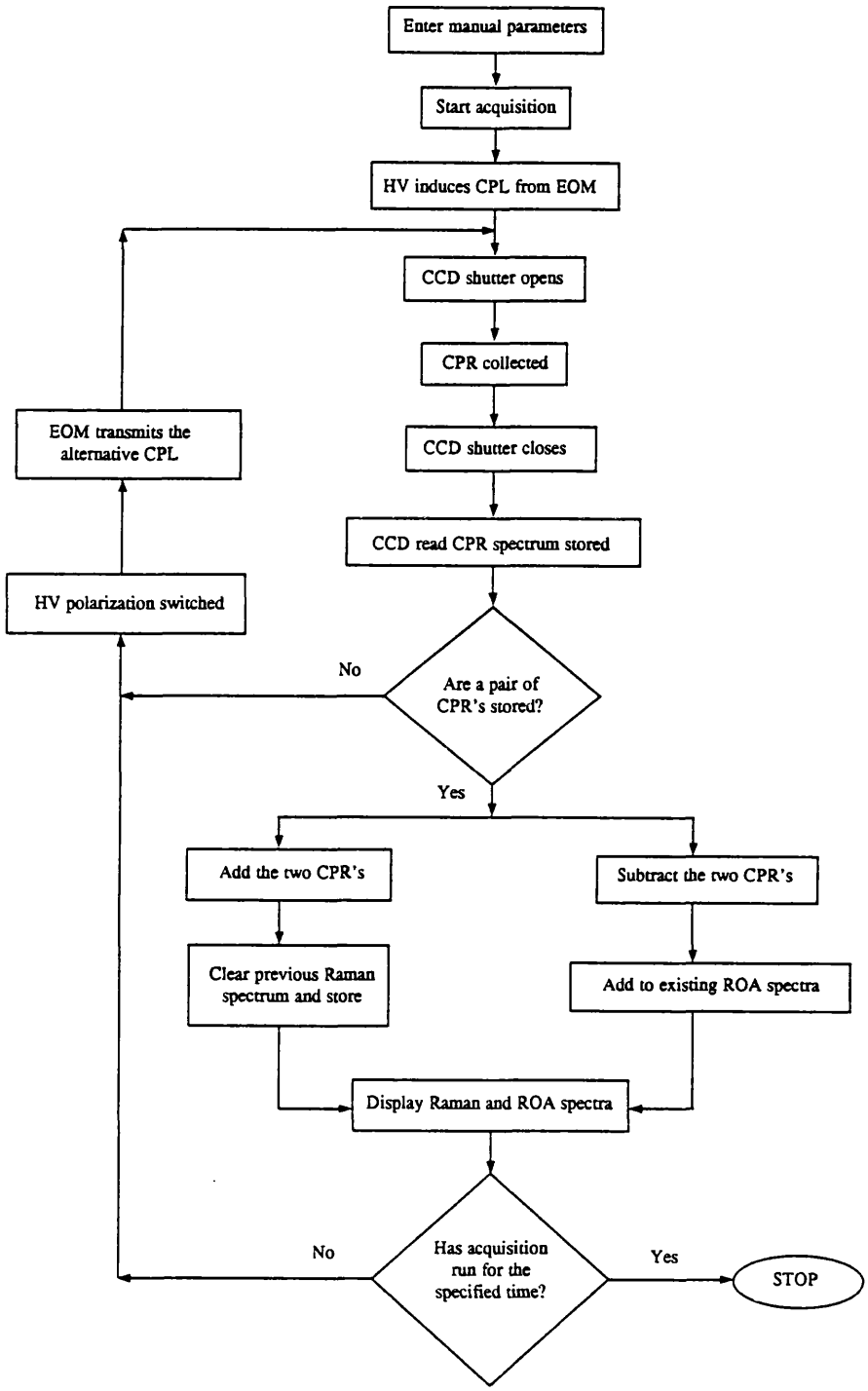


Figure 2.4: A flow diagram of the operation of the ROA spectrometer.

Abbreviations: HV - High Voltage produced by the differential amplifier. CPL - Circularly Polarized Light. CPR - Circularly Polarized Raman spectrum (the Raman spectrum emitted by a molecule under one of the two forms of circularly polarized light).

In cases like these the baseline is corrected useing a quadratic polynomical baseline correction function available within the Lab. Calc. software package. Cosmic ray strikes and artifacts are also removed from the final ROA spectrum using 'Zap' functions contained within Lab. Calc. Although these features usually appear in any one ROA spectrum, as more spectra are obtained from a particular sample, artifacts can be confidently determined.

2.2 THEORETICAL DEVELOPMENTS

2.2.1 Outline

This part of the introduction deals with the theory of the ROA phenomena. Although this is a complicated subject, rooted in electrodynamics and molecular quantum mechanics, it does have a place in this thesis for several reasons. Firstly, the theory illustrates the molecular properties that are responsible for ROA. Secondly, it will become apparent why the ROA strategy used at Glasgow is ideal for protein ROA. Thirdly, we will be able to outline how ab initio calculations are done (the results of such calculations are used in other parts of the thesis). Finally, the theoretical derivations could be described as 'elegant', and deducing and understanding such equations gives a great sense of satisfaction. The original approximate derivation, which is followed here, can be found in Barron (1982). The complete theoretical description of the ROA phenomena can be found in Hecht and Nafie (1991).

2.2.2 Basic mathematical and physical principles.

Over a century ago in 1852 G.G. Stokes proposed that a beam of polarized light could be described by four parameters, S_0, S_1, S_2 and S_3.

For the purposes of this discussion we only require the first parameter, which is related to the intensity I:

$$S_O = \tilde{E}_x \tilde{E}_x^* + \tilde{E}_y \tilde{E}_y^* = 2(\mu_o/\varepsilon_o)^{1/2} I \qquad 2.1$$

where **E** is the electric field vector of the light wave and μ_o and ε_o are the SI constants corresponding to the permittivity and permeability of free space. (The tildes indicate complex numbers.)

The interaction between a molecule and the electromagnetic field of a light wave is determined by the electric and magnetic properties of that molecule. These electric and magnetic characteristics are described by molecular multipole moments and those that dictate ROA are:

Electric dipole moment : $\mu_\alpha = \sum_i e_i r_{i\alpha}$

Magnetic dipole moment: $m_\alpha = \sum_i (e_i/2m_i)\varepsilon_{\alpha\beta\gamma} r_{i\beta} p_{i\gamma}$ 2.2

Magnetic dipole moment: $\Theta_{\alpha\beta} = (1/2)\sum_i e_i(3r_{i\alpha}r_{i\beta} - r_i^2\delta_{\alpha\beta})$

These equations are expressed in cartesian tensor notation, where particle i at position $\mathbf{r}_i$ has charge e_i, mass m_i and linear momentum $\mathbf{p}_i$. The Greek subscripts denote vector or tensor components and can be equal to x, y or z; a repeated suffix denotes summation over the three components (so that, for example, $a_\alpha b_\alpha \equiv \mathbf{a.b} = a_x b_x + a_y b_y + a_z b_z$). $\delta_{\alpha\beta}$ is the unit second-rank symmetric tensor, and $\varepsilon_{\alpha\beta\gamma}$ is the unit third-rank antisymmetric tensor.

The time dependent fields of the incident light wave induces the following oscillating multipole moments in the molecule:

$$\tilde{\mu}_\alpha^{(o)} = (\tilde{\alpha}_{\alpha\beta} + (i\omega/3c)n_\gamma \tilde{A}_{\alpha\gamma\beta} + (1/c)\varepsilon_{\delta\gamma\beta} n_\gamma \tilde{G}_{\alpha\delta} + \ldots)(\tilde{E}_\beta^{(o)})$$

$$\tilde{\Theta}_{\alpha\beta}^{(o)} = (\tilde{A}_{\gamma\alpha\beta}^* + \ldots)(\tilde{E}_\beta^{(o)}) \qquad 2.3$$

$$\tilde{m}_\alpha^{(o)} = (\tilde{G}_{\beta\alpha}^* + \ldots)(\tilde{E}_\beta^{(o)})$$

where $\tilde{\alpha}_{\alpha\beta}$, $\tilde{G}_{\alpha\beta}$ and $\tilde{A}_{\alpha\beta\gamma}$ are complex dynamic molecular property tensors (called the polarizability tensor, electric dipole-magnetic dipole optical activity tensor and the electric dipole-electric quadrupole optical activity tensor, respectively).

These complex tensors can be related to real tensors as follows:

$$\begin{aligned} \tilde{\alpha}_{\alpha\beta} &= \alpha_{\alpha\beta} - i\alpha'_{\alpha\beta} \\ \tilde{G}_{\alpha\beta} &= G_{\alpha\beta} - iG'_{\alpha\beta} \\ \tilde{A}_{\alpha\beta\gamma} &= A_{\alpha\beta\gamma} - iA'_{\alpha\beta\gamma} \end{aligned} \qquad 2.4$$

Using time-dependent perturbation theory the real dynamic molecular property tensors can be expressed quantum mechanically as follows:

$$\alpha_{\alpha\beta} = (2/\hbar)\sum_{j\neq n} (\omega_{jn}/\omega^2_{jn}-\omega^2)\ \mathrm{Re}(\langle n|\mu_\alpha|j\rangle\langle j|\mu_\beta|n\rangle)$$

$$\alpha'_{\alpha\beta} = -(2/\hbar)\sum_{j\neq n} (\omega_{jn}/\omega^2_{jn}-\omega^2)\ \mathrm{Im}(\langle n|\mu_\alpha|j\rangle\langle j|\mu_\beta|n\rangle)$$

$$G_{\alpha\beta} = (2/\hbar)\sum_{j\neq n} (\omega/\omega^2_{jn}-\omega^2)\ \mathrm{Re}(\langle n|\mu_\alpha|j\rangle\langle j|m_\beta|n\rangle) \qquad 2.5$$

$$G'_{\alpha\beta} = -(2/\hbar)\sum_{j\neq n} (\omega/\omega_{jn}^2-\omega\ )^2\mathrm{Im}(\langle n|\mu_\alpha|j\rangle\langle j|m_\beta|n\rangle)$$

$$A_{\alpha\beta\gamma} = (2/\hbar)\sum_{j\neq n} (\omega_{jn}/\omega^2_{jn}-\omega^2)\ \mathrm{Re}(\langle n|\mu_\alpha|j\rangle\langle j|\Theta_{\beta\gamma}|n\rangle)$$

$$A'_{\alpha\beta\gamma} = -(2/\hbar)\sum_{j\neq n} (\omega_{jn}/\omega^2_{jn}-\omega^2)\ \mathrm{Im}(\langle n|\mu_\alpha|j\rangle\langle j|\Theta_{\beta\gamma}|n\rangle)$$

2.2.3 Derivation of the ROA equations.

A scattered electric field vector, $\mathbf{E}^d$, will be related the properties of the scattering molecule and the properties of the incident light. Explicitly:

$$\tilde{E}^d_\alpha = (\omega^2\mu_o/4\pi R)\ \tilde{a}_{\alpha\beta}\tilde{E}^o_\beta \qquad 2.6$$

where $\tilde{a}_{\alpha\beta}$, the scattering tensor, describes the molecular properties and $\tilde{E}^o_\beta$ describes the incident light. The scattering tensor is related to the molecular multipole moments and so in turn is related to the complex dynamic molecular property tensors. In terms of the latter the scattering tensor is:

$$\begin{aligned}\tilde{a}_{\alpha\beta} = \tilde{\alpha}_{\alpha\beta} &+ (i\omega/3c)\ (n^d{}_\gamma\tilde{A}_{\alpha\gamma\beta} - n^d{}_\beta\tilde{A}_{\gamma\alpha\beta}{}^*) \\ &+ (1/c)\ (\varepsilon_{\delta\gamma\beta}n^d{}_\gamma\tilde{G}_{\alpha\delta} - \varepsilon_{\alpha\beta\gamma}n^d{}_\beta\tilde{G}_{\delta\gamma}{}^*)\end{aligned} \qquad 2.7$$

This analysis will tell us the dependence of the ROA signal on the scattering angle. So we must be able to convert the axis describing the scattered wave to the axis describing the incident wave. This is illustrated in Figure 2.5 where we see that the Stokes parameters of the scattered wave, which uses the coordinate system that it defines itself of x^d, y^d and z^d, must be converted to the coordinate system defined by the incident wave of x, y, and z.

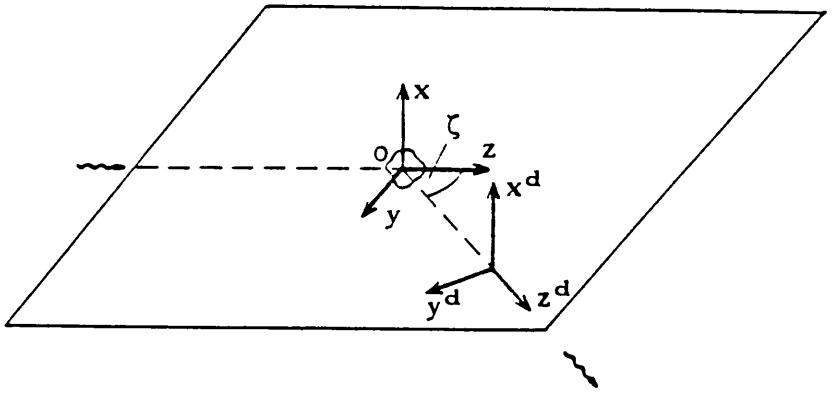


Figure 2.5: The coordinate system used to describe the incident wave (x, y and z) and the scattered wave (x^d, y^d and z^d). ζ is the scattering angle.

The exact equations that are used for the conversion are not reproduced here, but it suffices to say that both the scattering tensor and the Stokes parameters have to be converted to take account of the scattering angle ζ. The scattering angle-dependent scattering tensor is:

$$\tilde{a}_{\alpha\beta} = \tilde{\alpha}_{\alpha\beta} + (i\omega/3c)\ (\tilde{A}_{\alpha z\beta} - \tilde{A}_{\beta y\alpha}{}^{*}\sin\zeta - \tilde{A}_{\beta z\alpha}{}^{*}\cos\zeta) + (1/c)\ (\varepsilon_{\gamma z\beta}\tilde{G}_{\alpha\gamma} + \varepsilon_{\gamma y\alpha}\tilde{G}_{\beta\gamma}{}^{*}\sin\zeta + \varepsilon_{\gamma z\alpha}G_{\beta\gamma}{}^{*}\cos\zeta) \qquad 2.8$$

The angle dependent first Stokes parameter is:

$$S^{d}{}_{O} = \tilde{E}^{d}{}_{x}\tilde{E}^{d*}{}_{x} + \tilde{E}^{d}{}_{y}\tilde{E}^{d*}{}_{y}\cos^{2}\zeta + \tilde{E}^{d}{}_{z}\tilde{E}^{d*}{}_{z}\sin^{2}\zeta - (\tilde{E}^{d}{}_{y}\tilde{E}^{d*}{}_{z} + \tilde{E}^{d}{}_{z}\tilde{E}^{d*}{}_{y})\cos\zeta\sin\zeta. \qquad 2.9$$

The scattered electric field vector components are described in terms of the molecular properties by equation 2.6 and by combining 2.6 and 2.9 the first Stokes parameter of the scattered light (eventually) becomes:

$$\begin{aligned} S^{d}{}_{O} = {} & (1/2)(\omega^{2}\mu_{o}/4\pi R)^{2}\{(|\tilde{a}_{xx}|^{2}+|\tilde{a}_{xy}|^{2})S_{O}+(|\tilde{a}_{xx}|^{2}-|\tilde{a}_{xy}|^{2})S_{1} \\ & -2\mathrm{Re}(\tilde{a}_{xx}\tilde{a}_{xy}{}^{*})S_{2}-2\mathrm{Im}(\tilde{a}_{xx}\tilde{a}_{xy}{}^{*})S_{3}+[(|\tilde{a}_{xy}|^{2}+|\tilde{a}_{yy}|^{2})S_{O}+(|\tilde{a}_{xy}|^{2}-|\tilde{a}_{yy}|^{2})S_{1} \\ & -2\mathrm{Re}(\tilde{a}_{yx}\tilde{a}_{yy}{}^{*})S_{2}-2\mathrm{Im}(\tilde{a}_{yx}\tilde{a}_{yy}{}^{*})S_{3}]\cos^{2}\zeta+[(|\tilde{a}_{zx}|^{2}+|\tilde{a}_{zy}|^{2})S_{O} \\ & +(|\tilde{a}_{zx}|^{2}-|\tilde{a}_{zy}|^{2})S_{1}-2\mathrm{Re}(\tilde{a}_{zx}\tilde{a}_{zy}{}^{*})S_{2}-2\mathrm{Im}(\tilde{a}_{zx}\tilde{a}_{zy}{}^{*})S_{3}]\sin^{2}\zeta \\ & -2[\mathrm{Re}(\tilde{a}_{yx}\tilde{a}_{zx}{}^{*}+\tilde{a}_{yy}\tilde{a}_{zy}{}^{*})S_{O}+\mathrm{Re}(\tilde{a}_{yx}\tilde{a}_{zx}{}^{*}-\tilde{a}_{yy}\tilde{a}_{zy}{}^{*})S_{1} \\ & -\mathrm{Re}(\tilde{a}_{yx}\tilde{a}_{zy}{}^{*}+\tilde{a}_{zx}\tilde{a}_{yy}{}^{*})S_{2}-\mathrm{Im}(\tilde{a}_{yx}\tilde{a}_{zy}{}^{*}+\tilde{a}_{zx}\tilde{a}_{yy}{}^{*})S_{3}]\cos\zeta\sin\zeta\} \end{aligned} \qquad 2.10$$

The scattering factor definition in equation 2.8 is substituted into equation 2.10. The complex dynamic molecular tensors are then expanded and the whole equation is isotropically averaged, to take into account the fact that ROA is measured in a liquid, giving finally:

$$S^{d}_{O} = (1/2)(\omega^2\mu_o/4\pi R)^2\{(1/30)[(\alpha_{\alpha\alpha}\alpha_{\beta\beta}^{*}+7\alpha_{\alpha\beta}\alpha_{\alpha\beta}^{*})S_O+(3\alpha_{\alpha\alpha}\alpha_{\beta\beta}^{*}+\alpha_{\alpha\beta}\alpha_{\alpha\beta}^{*})S_1]$$

$$-(2ImS_3/30c)[(\omega/3)(1-3\cos\zeta)\varepsilon_{\alpha\beta\gamma}\alpha_{\delta\alpha}A_{\gamma\beta\delta}^{*}-\alpha_{\alpha\alpha}G'_{\beta\beta}{}^{*}-7\alpha_{\alpha\beta}G'_{\alpha\beta}{}^{*}$$

$$+5\cos\zeta(\alpha_{\alpha\beta}G'_{\alpha\beta}{}^{*}-\alpha_{\alpha\alpha}G'_{\beta\beta}{}^{*})]+(\cos^2\zeta/30)[(7\alpha_{\alpha\beta}\alpha_{\alpha\beta}^{*}+\alpha_{\alpha\alpha}\alpha_{\beta\beta}^{*})S_O$$

$$-(3\alpha_{\alpha\alpha}\alpha_{\beta\beta}^{*}+\alpha_{\alpha\beta}\alpha_{\alpha\beta}^{*})S_1-(2ImS_3/c)[(\omega/3)(1-3\cos\zeta)\varepsilon_{\alpha\beta\gamma}\alpha_{\delta\alpha}A_{\gamma\beta\delta}^{*}-\alpha_{\alpha\alpha}G'_{\beta\beta}{}^{*}$$

$$-7\alpha_{\alpha\beta}G'_{\alpha\beta}{}^{*}-5\cos\zeta(\alpha_{\alpha\alpha}G'_{\beta\beta}{}^{*}-\alpha_{\alpha\beta}G'_{\alpha\beta}{}^{*})]+(\sin^2\zeta/30)[(6\alpha_{\alpha\beta}\alpha_{\alpha\beta}^{*}-2\alpha_{\alpha\alpha}\alpha_{\beta\beta}^{*})S_O$$

$$-(2ImS_3/c)((-\omega/3)(2+\cos\zeta)\varepsilon_{\alpha\beta\gamma}\alpha_{\delta\alpha}A_{\gamma\beta\delta}+2(\alpha_{\alpha\alpha}G'_{\beta\beta}{}^{*}-3\alpha_{\alpha\beta}G'_{\alpha\beta}{}^{*})]$$

$$+(2ImS_3\sin^2\zeta\cos\zeta/30c)[\omega\varepsilon_{\alpha\beta\gamma}\alpha_{\delta\alpha}A_{\gamma\beta\delta}^{*}-5(\alpha_{\alpha\beta}G'_{\alpha\beta}{}^{*}-\alpha_{\alpha\alpha}G'_{\beta\beta}{}^{*})]\} \qquad 2.11$$

We now have an equation that tells us the intensity of a beam of light scattered at any particular angle from a molecule, and how that scattered intensity depends on the four Stokes parameters of the incident light. This has to be converted to give an expression of the quantity that is measured during the ROA experiment. That quantity is the Circular Intensity Difference (CID), Δ, and is defined by:

$$\Delta = (S^{dR}_{O}-S^{dL}_{O})/(S^{dR}_{O}+S^{dL}_{O}) \qquad 2.12$$

where S^{dR}_{O} and S^{dL}_{O} are the first Stokes parameters of light scattered by a molecule under incident circularly right and left polarized light respectively. (Although in reality the numerator and denominator of this equation are rarely compared, this equations gives a very convenient way of simplifing equation 2.11.) S^{dR}_{O} and S^{dL}_{O} are calculated by entering the appropriate incident Stokes parameters for circularly right and left polarized light in equation 2.11. The CID equation finally becomes:

$$\Delta = \{2[(1+\cos^2\zeta)[(2/3)\beta(A)^2(1-3\cos\zeta)+(2/3)7\beta(G')^2+45\alpha G'$$

$$-(10/3)\cos\zeta(\beta(G')^2-9\alpha G')]+\sin^2\zeta[(-4/3)\beta(A)^2(1+2\cos\zeta)+4\beta(G')^2$$

$$+(20/3)\cos\zeta(\beta(G')^2-9\alpha G')]\}/\{c[(1+\cos^2\zeta)[(14/3)\beta(\alpha)^2+30\alpha^2]$$

$$+(4\sin^2\zeta)\beta(\alpha)^2]\} \qquad 2.13$$

where:

$$\alpha = (1/3)\alpha_{\alpha\alpha}$$
$$G' = (1/3)G'_{\alpha\alpha} \qquad 2.14a$$

are called the isotropic invariants and:

$$\beta(\alpha)^2 = (1/2)(3\alpha_{\alpha\beta}\alpha_{\alpha\beta} - \alpha_{\alpha\alpha}\alpha_{\beta\beta})$$
$$\beta(G')^2 = (1/2)(3\alpha_{\alpha\beta}G'_{\alpha\beta} - \alpha_{\alpha\alpha}G'_{\beta\beta}) \qquad 2.14b$$
$$\beta(A)^2 = (1/2)\omega\alpha_{\alpha\beta}A_{\gamma\delta\beta}\varepsilon_{\alpha\gamma\delta}$$

are called the anisotropic invariants of the corresponding tensors and tensor products.

2.2.4 The implications of the theoretical analysis.

These equations lead to several very important deductions about the ROA experiment.

By examining the definition of the invariants responsible for ROA we can see how the ROA signal is generated. All the invariants in the numerator are products of the polarizibility tensor and one of the two optical activity tensors. This contrasts with optical rotation and circular dichroism experiments in fluids where the signal is generated neither by this interference (the product of the two tensors) or by the electric dipole-electric quadrupole optical activity tensor ($A_{\alpha\beta\gamma}$), but by the isotropic part of the electric dipole-magnetic dipole optical activity tensor ($G'_{\alpha\beta}$) only.

Secondly, in deducing the final CID equation it becomes apparent that the S_0 terms (that is terms dependent on the first Stokes parameter of the incident beam) vanish in the numerator. The S_3 terms, which are the only ones that appear in the numerator, and so generate the ROA signal,

are three orders of magnitude smaller than the S_0 terms. Consequently, any small imperfections in the polarization modulation system that results in different circularly left and right intensities will produce very large artifacts.

Thirdly, we can see that equation 2.10 has no term that is dependent upon S_2. Therefore, any change in the S_2 parameter during the polarization modulation will not affect the ROA result. In addition to this we see that the S_1 parameter, which should be zero in pure circularly polarized light, regulates large molecular polarizablility tensor terms. Therefore, any deviations from pure circular polarization will produce large artifacts.

It can be seen from the CID equation 2.13 that the contributions of $\alpha G'$, $\beta(A)^2$ and $\beta(G')^2$ to the numerator depend on the scattering angle. These contributions are shown by the graph in Figure 2.6, which demonstrates that the ROA signal is strongest when the scattering angle is 180°; i.e. a backscattering strategy is the optimum one for an ROA experiment.

2.2.5 Ab initio ROA calculations.

In some parts of this thesis peptide ab initio ROA calculations are used to interpret experimental peptide ROA. Ab initio calculations use the above equations to produce a theoretical ROA spectrum and the advantage of such calculations is that they give deeper insights into the vibrations involved in a particular ROA band. Such calculations were carried out on alanine, and experimental ROA spectra were published together with their theoretical counterpart in 1991 (Barron et al., 1991 and 1992a; Gargaro, 1991). Although ab initio calculations have only been done so far on small molecules, such as alanine, their results can be useful in the analysis of the ROA of larger molecules, such as alanyl peptides (Ford et al., 1994).

Some of the latest ab inito calculations on small peptides show how ROA changes with Φ and Ψ (Polavarapu and Deng, 1994). The basic ab initio ROA calculation method has been pioneered by Polavarapu (1990), following work by Amos (1982).

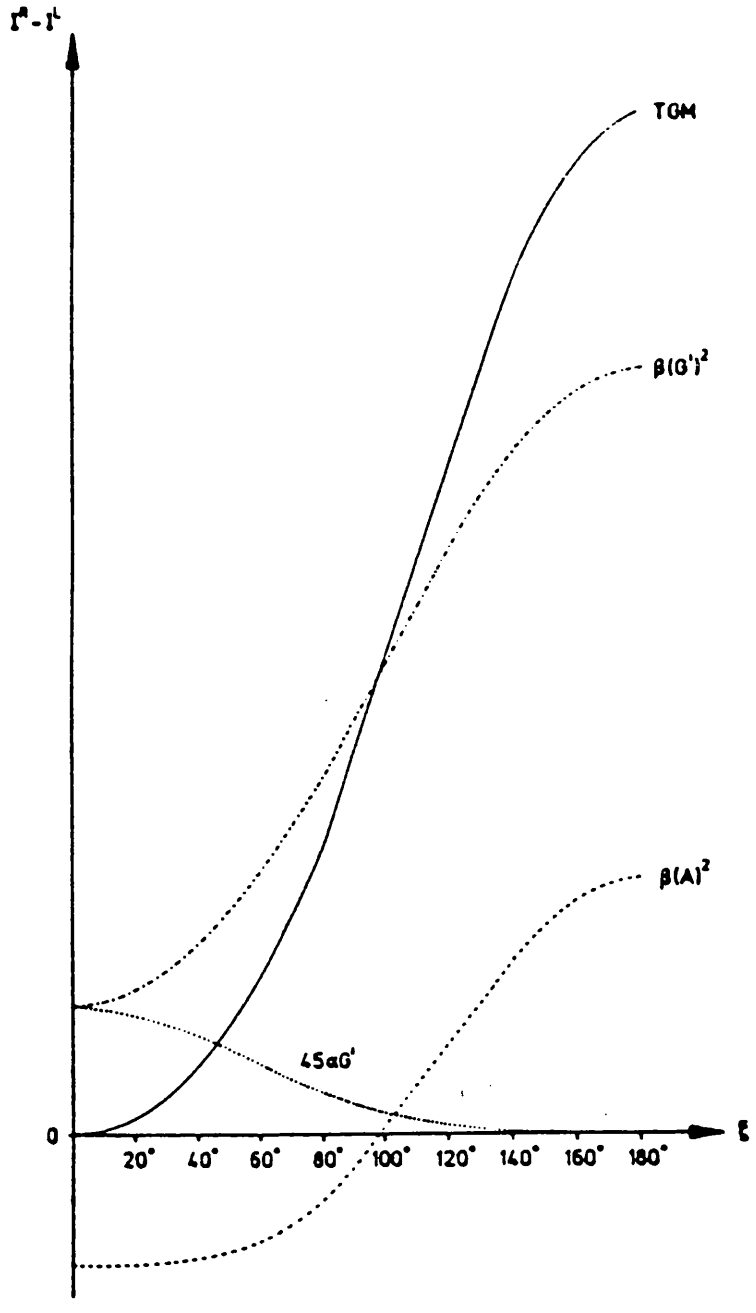


Figure 2.6: The dependence of $\alpha G'$, $\beta(A)^2$ and $\beta(G')^2$ on scattering angle. TGM represents the two group model which predicts that $\alpha G'=0$ and $\beta(A)^2=\beta(G')^2$. (Adapted from Hecht and Barron, 1990)

The CID equations described in section 2.2.3 were only developed to encompass Rayleigh optical activity (because only that was required to demostrate the backscattering advantage). When we turn to ROA calculations the Rayleigh CID is converted into the Raman CID by making the following substitutions (using what is called Placzek's approximation):

$$\alpha_{\alpha\beta} \rightarrow \langle \upsilon_m | \alpha_{\alpha\beta}(Q) | \upsilon_n \rangle$$

$$G'_{\alpha\beta} \rightarrow \langle \upsilon_m | G'_{\alpha\beta}(Q) | \upsilon_n \rangle \qquad 2.15$$

$$A_{\alpha\beta\gamma} \rightarrow \langle \upsilon_m | A_{\alpha\beta\gamma}(Q) | \upsilon_n \rangle$$

where υ_m and υ_n represent the initial and final vibrational wavefunctions and Q is the normal vibrational coordinate. This equation is expanded as a Taylor series in each normal mode coordinate, Q_p, as follows:

$$\langle \upsilon_m | \alpha_{\alpha\beta} | \upsilon_n \rangle = (\alpha_{\alpha\beta})_o \delta_{\upsilon_m \upsilon_n} + \sum_p (\partial \alpha_{\alpha\beta} / \partial Q_p)_o \langle \upsilon_m | Q_p | \upsilon_n \rangle + (1/2) \sum_{p,q} (\partial^2 \alpha_{\alpha\beta} / \partial Q_p \partial Q_q)_o \langle \upsilon_m | Q_p Q_q | \upsilon_n \rangle + \ldots \qquad 2.16$$

where the first term describes Rayleigh scattering, the second term describes fundamental vibrational Raman scattering and the third term describes overtones and combinations. Subscript zero refers to the equilbrium nuclear configuration. Similar expressions can be derived for $\langle \upsilon_m | G'_{\alpha\beta} | \upsilon_n \rangle$ and $\langle \upsilon_m | A_{\alpha\beta\gamma} | \upsilon_n \rangle$.

We saw that the Rayleigh CID equations used above were determined by the polarizability - polarizability tensor products, $\beta(\alpha)^2$, and the two polarizability-optical activity tensor products, $\beta(G')^2$ and $\beta(A)^2$. By using the Rayleigh definitions of these in terms of $\alpha_{\alpha\beta}$, $G'_{\alpha\beta}$ and $A_{\alpha\beta\gamma}$ and expanding them to Raman definitons (with the Placzeck approximation and

a two term Taylor series) we find that the central expressions governing vibrational Raman optical activity are:

$$(\hbar/2\omega_p)(\partial\alpha_{\alpha\beta}/\partial Q_p)_o(\partial\alpha_{\alpha\beta}/\partial Q_p)_o$$

$$(\hbar/2\omega_p)(\partial\alpha_{\alpha\beta}/\partial Q_p)_o(\partial G'_{\alpha\beta}/\partial Q_p)_o \qquad 2.17$$

$$(\hbar/2\omega_p)(\partial\alpha_{\alpha\beta}/\partial Q_p)_o(\partial A_{\alpha\gamma\beta}/\partial Q_p)_o$$

In order for the ab intio calculations to be carried out two approximations are made. The first is that the ROA experiment is carried out at transparent frequencies; the energy of the exciting laser light is far smaller the energy difference between the ground state and first exited electronic state. This converts the polarizability tensor to an approximate static form:

$$\alpha_{\alpha\beta} = 2\sum_{j=n}(1/W_{jn})\ \mathrm{Re}(\langle n|\mu_\alpha|j\rangle\langle j|\mu_\beta|n\rangle) \qquad 2.18$$

where W_{jn} is the energy difference between the energy of the jth virtual intermediate state and the energy of the initial state n. The optical activity tensors reduce to similar expressions. The second approximation is that the summation term can be accounted for by considering the effect of a static electric field on a wavefunction using time independent perturbation theory. This gives:

$$\alpha_{\alpha\beta} = 2\langle\Psi^{(O)}{}_n|\mu_\alpha|\Psi^{(1)}{}_n(E_\beta)\rangle \qquad 2.19$$

Similar expressions are derived for the optical activity tensors. By expressing the wavefunctions, Ψ, in terms of molecular orbitals it is possible to calculate the polarizability and optical activity tensors.

The derivatives that the central ROA expressions require, $(\partial\alpha_{\alpha\beta}/\partial Q_p)_o$, $(\partial G'_{\alpha\beta}/\partial Q_p)_o$ and $(\partial A_{\alpha\beta\gamma}/dQ_p)_o$, are calculated numerically by evaluating

$\alpha_{\alpha\beta}$, $(1/\omega)G'_{\alpha\beta}$ and $A_{\alpha\beta\gamma}$ at the equilibrium geometry and at geometries displaced by 0.005 $\mathring{A}$ along each atomic co-ordinate.

Certain problems exist with these calculations and one difficulty is the gauge origin dependence of the $G'_{\alpha\beta}$ calculation. However, recently Helgaker and coworkers (1994) have developed a series of vibrational ROA calculations using London atomic orbitals which not only yield gauge independent atomic orbitals, but also provide a set of equations which converge faster.

Chapter 3

Alanyl Peptides

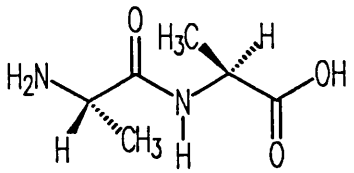
3.1 INTRODUCTION

The understanding of protein ROA spectra can be approached from two directions. The first is the accumulation of a large number of protein ROA spectra and the deduction of spectral-structural relationships, and the second is the step-by-step examination of increasingly large model systems. This particular section concerns the latter pathway, where a series of simple alanyl oligopeptides were studied by ROA.

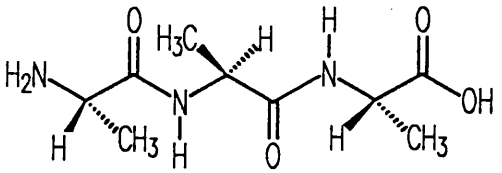
There were four reasons for choosing to study a series of alanyl peptides. Firstly alanine is the simplest chiral amino acid. Secondly, alanine and it's peptides have already been studied by IR, NMR, VCD, Raman, ROA, normal mode analysis and molecular mechanics calculations. Thirdly, the ROA of alanine is well understood via ab initio calculations (Barron et al., 1991 and 1992b; Gargaro, 1991), and finally the alanyl peptides are readily available from commercial sources.

Initially the study of the alanyl peptides consisted of Ala_2 and Ala_3 in H_2O and Ala_4 in aqueous HCl and these results have already been presented in the doctorial thesis of Z.Q. Wen (1992). However, it became apparent that a more extensive study of this simple system would prove more rewarding. Consequently we studied Ala_2, Ala_3, Ala_4 and Ala_5 (shown in Figure 3.1) in H_2O and D_2O, under a variety of pH conditions.

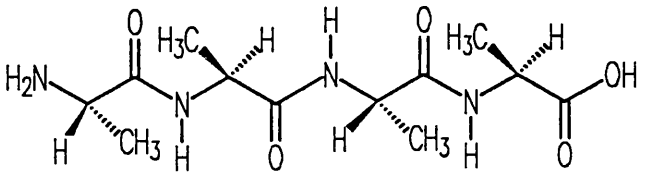
Ala_2



Ala_3



Ala_4



Ala_5

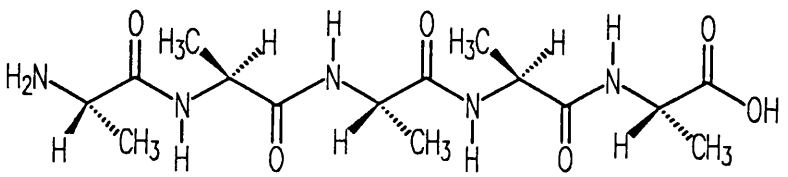


Figure 3.1: Structural formulae of the alanyl peptides.

3.2 EXPERIMENTAL

The Ala_2, Ala_3, Ala_4, Ala_5 samples, DCl and NaOD were bought from Sigma Chemical Co. and used without further purification. The same preparation protocol was used for the peptide samples as for the protein samples (see section 6.2.1).

For the deuterated studies of Ala_2 completely deuterated Ala_2 was used. This was prepared by dissolving Ala_2 in D_2O, and recovering it by lyophilization. The concentration of the other alanyl peptides was sufficently low to ensure that less than 2% of the peptide molecules had more than 1 proton unexchanged. (Concentrated HCl, dissolved in D_2O, was used to acidify the Ala_4 and Ala_5/DCl solutions, but again the

proton:deuterium ratio was sufficiently low to ensure the same condition as above.) The alanyl peptide concentrations and solutions are given in Table 3.1.

Table 3.1: The concentrations and conditions of the alanyl peptide solutions. Acid or alkali concentration in parenthesis. * indicates that the sample was prepared by Dr. Z.Q.Wen and no records are available to determine these concentrations.

	NaOD	NaOH	D_2O	H_2O	DCl	HCl
Ala_2	3.3M(5M)	*	3.1M	*	3.5M(5M)	*
Ala_3	.67M(1M)	.61M(1M)	.53M	*	.61M(1M)	.58M(1M)
Ala_4	.31M(1M)	.27M(1M)	-	.05M	.20M(.2M)	*
Ala_5	-	-	-	-	.08M(.1M)	.09M(.1M)

3.3 ASSIGNMENTS OF THE ALANYL PEPTIDE RAMAN AND ROA SPECTRA.

3.3.1 Overview

The quality of the Ala_2 and Ala_3 spectra are very good allowing small ROA features to be seen. However as the series increases peptide solubility decreases and the quality of the ROA falls. So, in the larger peptides only the largest ROA features can be seen. The assignments above 1200 cm^{-1} are relatively reliable, because the vibrations are localized and good normal mode calculations already exist (Diem et al., 1992; Qian et al., 1991): below 1200 cm^{-1} the vibrational modes are made up of several different contributions from various localized groups within the molecules and so are somewhat more difficult to assign.

One of the advantages of ROA over conventional Raman is that the comparision of experimental spectra and ab initio calculations are made easier by the extra 'chiral' dimension. The same is true for the comparison of two separate experimental ROA spectrum; the assignment of the dialanine spectrum is greatly helped by comparing similar ROA features in the alanine and Ala_2 spectra. A similar advantage is gained as we try to assign the longer alanyl peptides.

The final point before we start the analysis is to define some of the terms that we will use. These terms have been used for brevity, but it is important to define them as they can easily be confused with some other terms in chemistry. The solvents used were HCl, DCl, H_2O, D_2O, NaOH and NaOD. HCl and DCl will collectively be called acidic; H_2O and D_2O neutral; and NaOH and NaOD alkaline. HCl, H_2O and NaOH will collectively be called hydrated solvents. Similarly DCl, D_2O and NaOD will be termed deuterated solvents.

3.3.2 Ala_2

The Raman and ROA spectra of Ala_2 in HCl, DCl, H_2O, D_2O, NaOH and NaOD are shown in Figures 3.2 and 3.3. The low frequency Raman and ROA spectra of Ala_2 in H_2O and D_2O are shown in Figure 3.4.

1750-1450 cm^{-1}

The Raman and ROA in the region between 1750 and 1500 cm^{-1} give themselves over to a nice clear explanation. The frequencies and ultimate assignments are shown in Table 3.2, with the discussion following on.

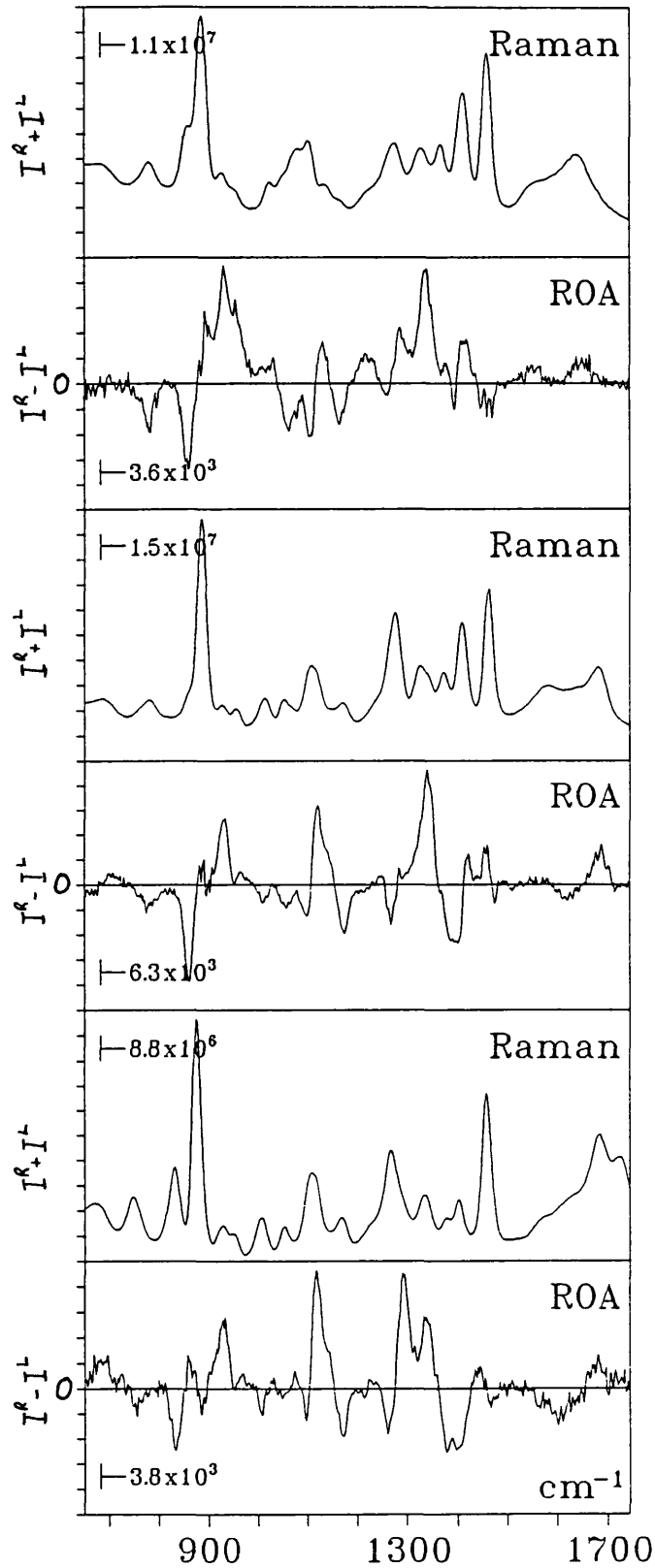


Figure 3.2: Ala_2 in hydrated solvents: NaOH (top), H_2O (middle), HCl (bottom).

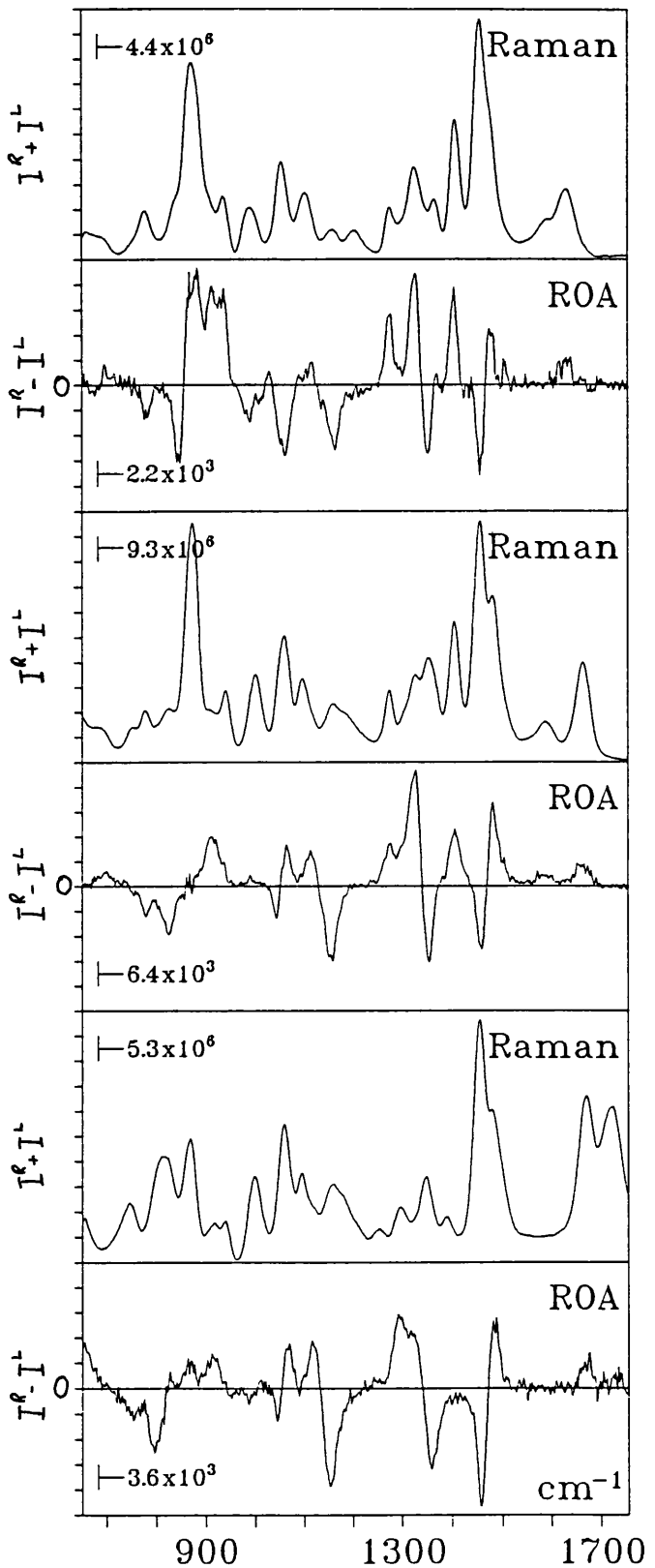


Figure 3.3: Ala_2 in deuterated solvents: 5M NaOD (top), D_2O (middle), 5M DCl (bottom).

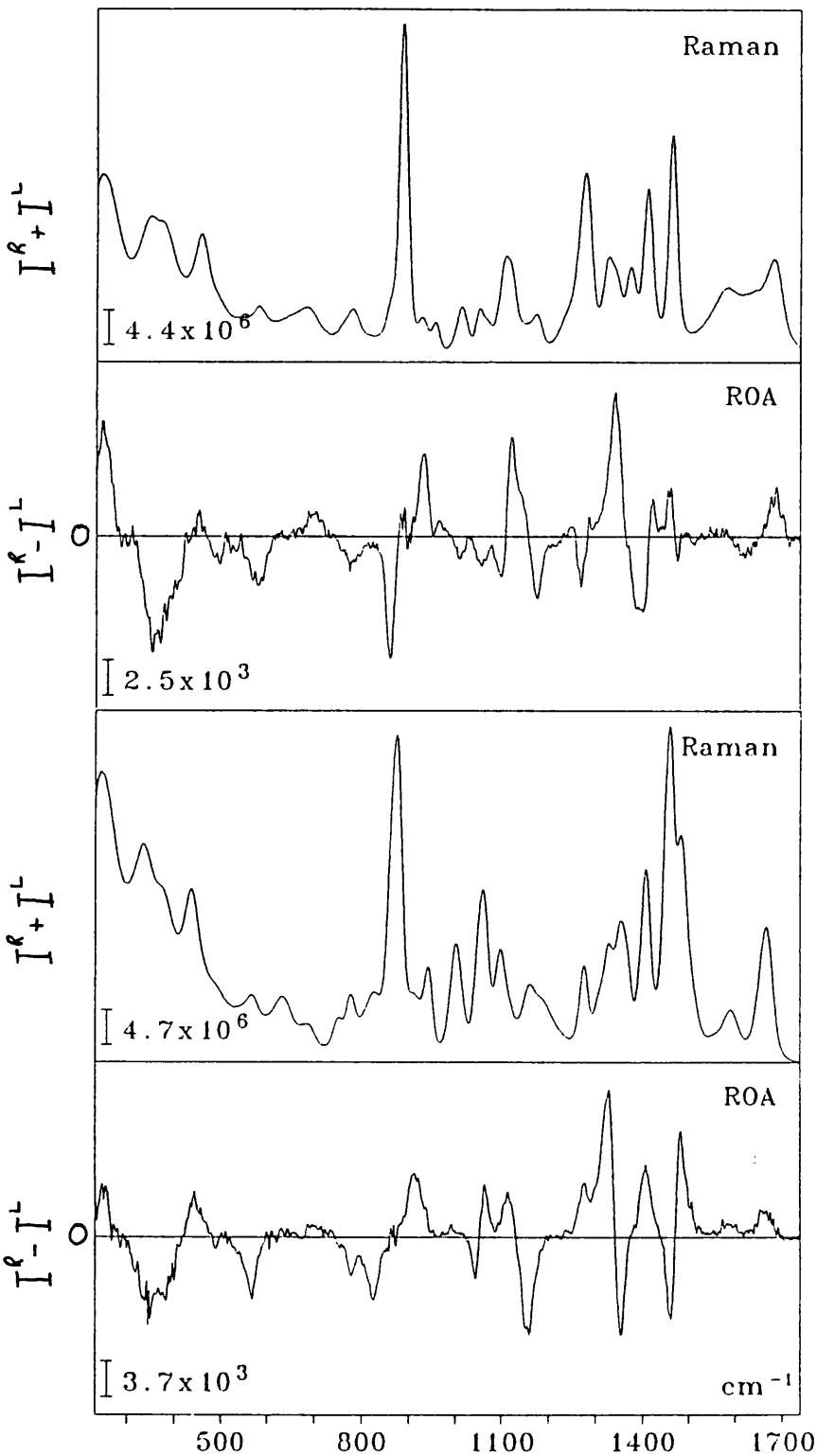


Figure 3.4: Raman and ROA spectra (between 250 and 1750 cm^{-1}) for Ala_2 in H_2O (top) and D_2O (bottom).

Table 3.2: The frequencies and assignments of the Raman and ROA signals of the amide I (and I') region of Ala_2. The figures in parenthesis indicate the approximate relative intensity of the ROA associated with the band. (st; stretch: asym; asymmetric: def; deformation.)

HCl	DCl	H_2O	D_2O	NaOH	NaOD	Assignment
1720(0)	1720(0)					$CO_2H(D)$[a]
1675(+1)		1665(+1)				Amide I[b]
	1665(+1)		1660(+1)			Amide I'[b]
				1630(+1)		Amide I
					1625(+1)	Amide I'
1620(-1.5)		1620(-0.5)[c]				NH_3^+ asym def[d]
			1580(+0.5)		1580(0)	CO_2^- asym st[b]
1570(-0.5)		1570(0)		1575(0)[c]		CO_2^- asym st /amide II[b,e]
				1550(+1)		Amide II/NH_2 def

a = Williams and Fleming, 1989
b = Oboodi et al., 1984
c = these are bands that cannot be distinguished from the surrounding bands, but are present. (The frequencies are not very accurate.)
d = Kettle et al., 1990
e = Diem et al., 1992

It is clear from these results that under acidic and neutral conditions the amide I remains the same, but in alkali solvents the band falls by almost 40 cm^{-1}. This also occurs with the amide I' band which is 5-10 cm^{-1} lower than the amide I. This observation was interesting because previous studies on Ala_3 (Lee et al., 1989) had proposed that this 40 cm^{-1} drop demonstrated a significant conformational change in Ala_3. This frequency drop is discussed in more detail in section 3.3.3 but it appears as if this frequency change arises from a change in the vibrational normal mode.

In HCl and H_2O the CO_2^- asymmetric stretch and the amide II vibrations occur together, but in NaOH they separate. Although this is not

immediately clear the D_2O and NaOD results show that the CO_2^- asymmetric stretch remains stationary between neutral and alkali pH whereas the amide II' band is lowered in alkali conditions (see below). It is reasonable to assume that the same will happen in hydrated solvents and so it appears as if the 1550 cm^{-1} band in the Raman spectrum of NaOH is the lowered amide II vibration (possibly with NH_2 deformation). (An alterative hypothesis that ignores the amide II' result is that the amide II vibration remains at 1570 cm^{-1} in NaOH with the CO_2^- asymmetric stretch, and the 1550 cm^{-1} band is caused by NH_2 deformations.)

The asymmetric methyl deformation bands occur at ~1460 cm^{-1} in all six spectra and the frequency remains virtually constant. In the hydrated solvents the ROA is consistently a positive/negative couplet, although the intensity of the high and low frequency components seems to alter with pH, as does the equivalent ROA couplet in alanine (Barron et al., 1991 and 1992b; Gargaro, 1991), indicating possible interactions with the end groups. The couplet itself arises from either the splitting of the degenerate asymmetric methyl deformations by other vibrations or the different environments of the methyl groups, or both (Barron et al., 1991; Ford et al., 1994).

The situation is more complicated in the deuterated solutions. Here the asymmetric methyl stretch has a high frequency shoulder at ~1465 cm^{-1} which has been the assigned to the amide II' stretch: these two bands carry a sharp negative/positive ROA couplet in all the deuterated solvent spectra. Although the amide II vibration (in hydrated solvents) is weak, the proximity of the amide II' vibration to the strong asymmetric methyl stretch allows it to borrow intensity from it (Oboodi et al., 1984). This idea is reinforced by the sharp ROA couplet that embraces both the asymmetric methyl stretches and the amide II' vibrations. As was noted

above, the Raman amide II' frequency falls in NaOD, becoming only a slight shoulder on the methyl stretch band.

1400-1200 cm^{-1}

The Raman and ROA spectra between 1200 and 1420 cm^{-1} for all the alanyl peptides in hydrated and deuterated solvents are shown in Figures 3.5 and 3.6.

The first bands in this region are the 1370 and 1410 cm^{-1} Raman bands in neutral and alkali solutions, which are assigned to the symmetric methyl deformation and the carboxylate symmetric stretch, respectively (Oboodi et al., 1984). (In D_2O the methyl symmetric deformations form a shoulder of the 1350 cm^{-1} C-H deformation bands.) Under acidic conditions the carboxylate stretch band vanishes, leaving behind a small residual band, and the frequency and intensity of the symmetric methyl deformation alters.

In H_2O and HCl we see that the ROA generated by the methyl group (a negative peak) and the carboxylate group (a negative/positive couplet) coincides, indicating that vibrational interactions occur between the two groups. In deuterated solvents the positive ROA band caused by the CO_2^- stretch collapes completely upon deuteration.

The Raman spectra under alkali and neutral conditions remains similar, but the ROA of the symmetric methyl deformation appears to alter as pH rises. In hydrated solvents the negative/positive couplet of the carboxylate stretch remains, but the distinctive negative ROA of the symmetric methyl deformation vanishes in NaOH. In deuterated solvents the situation is not so clear cut. These features can be explained by a NH_3^+ involvement in the vibrations of this region at neutral pH. The NH_3^+ symmetric deformations

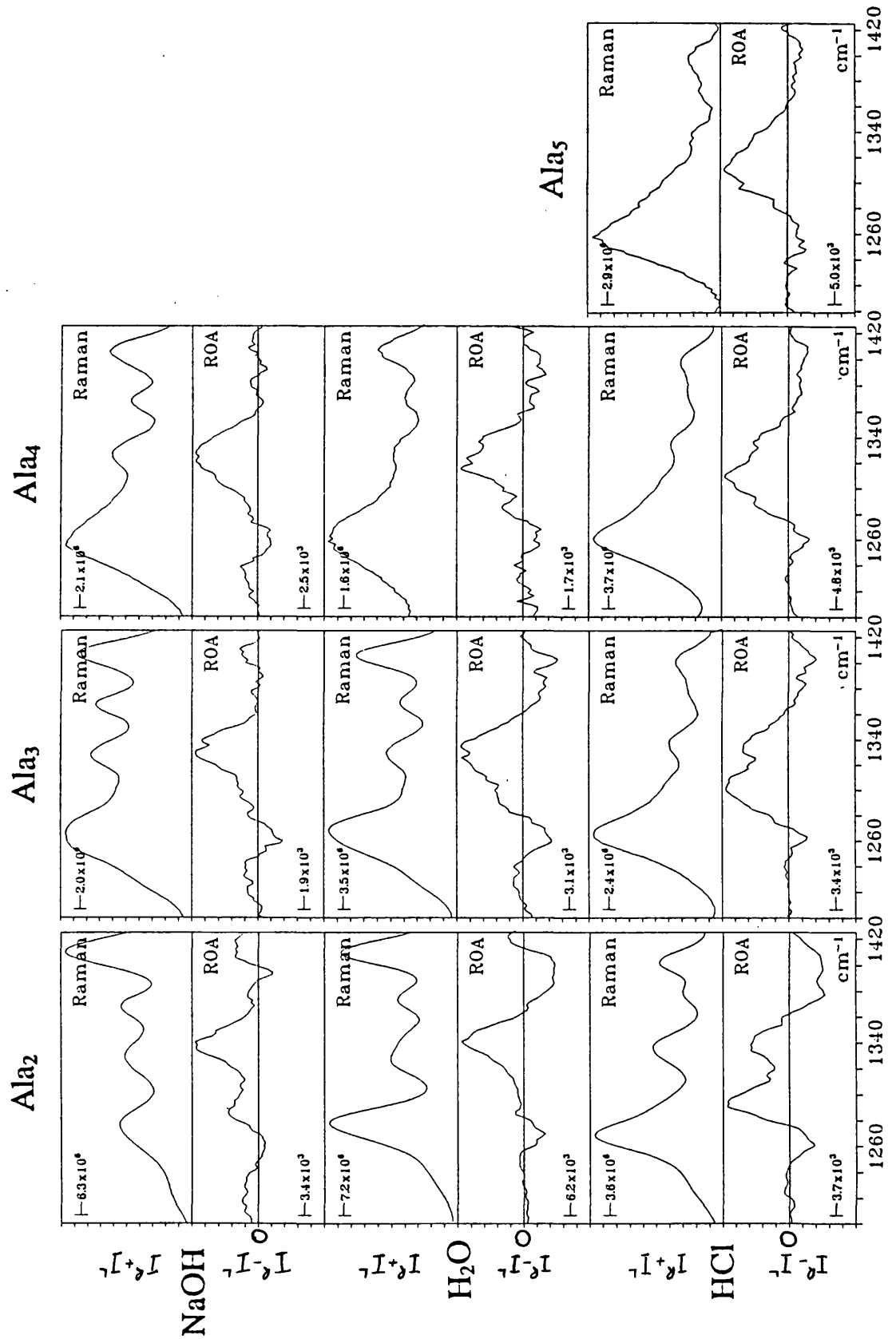


Figure 3.5: The amide III region (1200-1420 cm^{-1}) of the alanyl peptides.

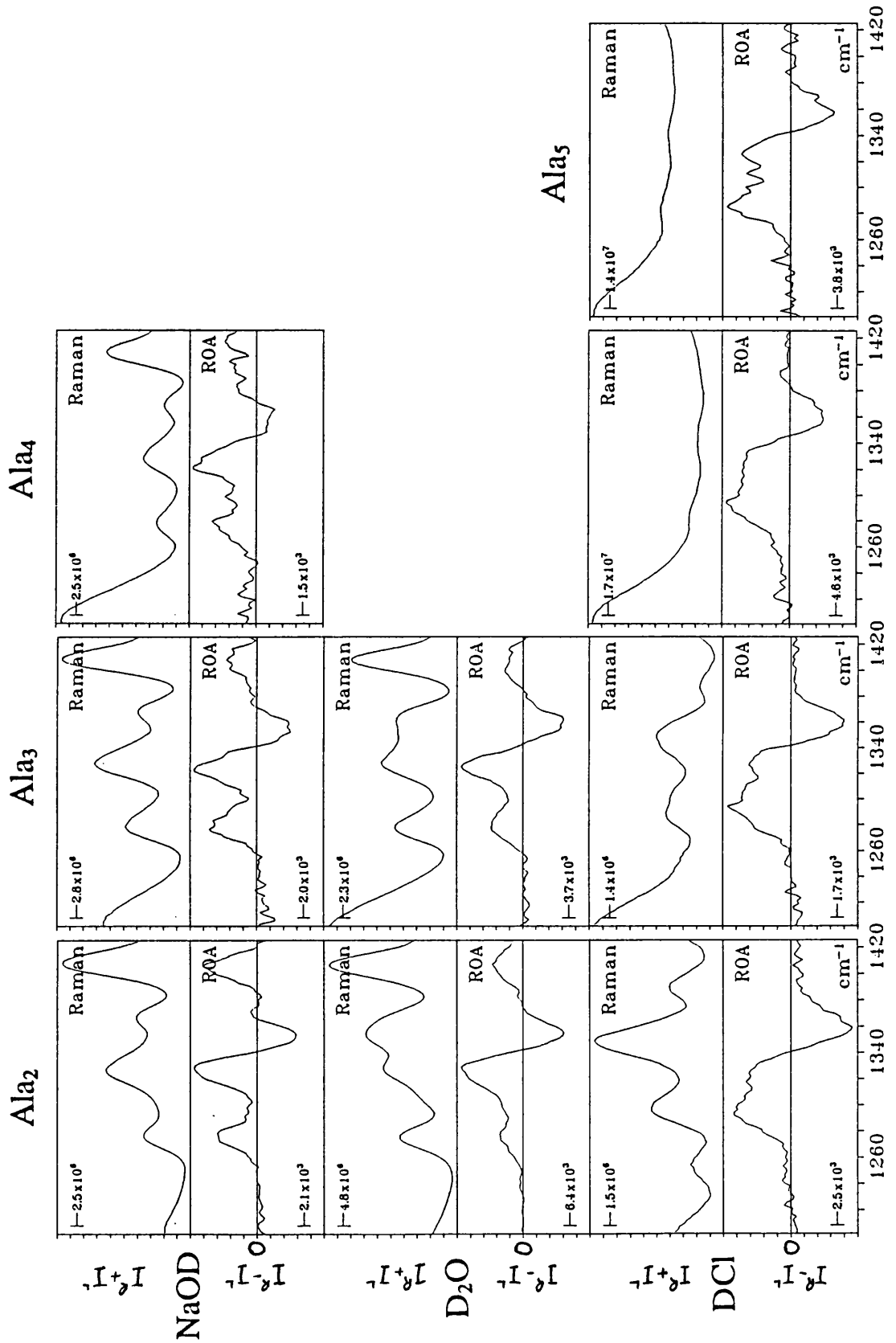


Figure 3.6: The 1200-1400 cm^{-1} region of the alanyl peptides in deuterated solvents.

appear in this region and make a contribution to the ROA of alanine in this region (Gargaro, 1991). The NH_3^+ symmetric deformations will alter as the pH rises, in contrast to the ND_3^+ vibrations which will come elsewhere, 1210 cm^{-1} (Kettle et al., 1990), and will have no effect on this region. This concurs nicely with the larger degree of change we see between H_2O and NaOH than between D_2O and NaOD solutions.

We now come to the amide III region where C-H and N-H deformations mix, producing vibrations characteristic of polypeptide structure (Krimm and Bandekar, 1986). This region is critically important for many peptide and protein studies.

'Monomeric' alanine has two C-H deformations at 1307 and 1355 cm^{-1} (Roberts et al., 1988) which start to mix in the higher peptides, and so for convenience we label these monomeric vibrations C-H^I and C-H^{II}, respectively. In Ala_2 the subscripts N and C refer to the deformations originating on the amino and carboxylate terminal respectively.

If alanine has two methine deformations Ala_2 would be expected to have four. However, in hydrated Ala_2 these four deformations mix. In D_2O the amide link nitrogen is deuterated: this lowers the N-H deformation (which occurs about 1336 cm^{-1} in the hydrated Ala_2) to 1005 cm^{-1} (Oboobi et al., 1984), and prevents the C-H mixing across the amide link. Consequently Ala_2 in D_2O has four unperturbed C-H deformations at 1276cm^{-1} , 1305cm^{-1} , 1329 cm^{-1} , 1355 cm^{-1} (Oboobi et al., 1984) which corresponds to the C_C-H^I , C_N-H^I, C_C-H^{II} and C_N- H^{II} vibrations respectively.

Diem and co-workers (1992) have combined normal mode calculations with IR and VCD data to provide qualitative estimates for the mixing of the four C-H and N-H deformations in Ala_2 in H_2O. They showed that

within the 1250-1360 cm^{-1} region the C-H and N-H vibrations couple to produce three mixed amide III bands (designated amide III^1 , amide III^2 and amide III^3 in order of ascending frequency). The contributions to these vibrations are shown in Table 3.3.

Table 3.3: The amide III vibrations of Ala_2.

Vibration	Frequency	Contributions
Amide III^1	(1281 cm^{-1})	N-H ipd, C-N st, C_C-H^{II}, C_N-H^{II}
Amide III^2	(1325 cm^{-1})	C_C-H^{II}, C_N-H^{II} (with a "negligible" N-H deformation)
Amide III^3	(1340 cm^{-1})	N-H ipd, C_N-H^{II}

The C_C-H^I and C_N-H^I vibrations do not mix and occur at 1266 and 1302 cm^{-1} respectively.

The Ala_2 analyses in D_2O and H_2O are the easiest because we already have the normal mode analysis and the molecule is only 'one step' up from the well understood Ala molecule. In H_2O we see only a broad band between 1250 and 1300 cm^{-1}, but Diem has shown by deconvolution that there are three vibrations here, the C_C-H^I at 1266 cm^{-1}, the amide III^1 at 1281 cm^{-1} and the C_N-H^I at 1302 cm^{-1} (the latter is very weak). The frequency of the negative ROA band in this region (1270 cm^{-1}) appears to indicate that it originates from the C_C-H^I deformation. However, it would be surprising if the highly delocalized amide III^1 mode gave no ROA and in fact the Ala_2/D_2O results shows a positive ROA band in this region, demonstating that either the C_C-H^I vibration is not as pure as originally thought, or that the amide III^1 ROA has a larger effect. The C_N-H^I deformation gives no perceptible ROA. The amide III^3 vibration (~1340 cm^{-1}) gives very strong positive ROA. Although the ROA of the amide III^2 band (1322 cm^{-1}) is also positive it is merely a shoulder of the amide III^3 ROA band.

In D_2O we see that the C_C-H^I (1265 cm^{-1}) and C_N-H^I (1305 cm^{-1}) deformations both give positive ROA. The C_C-H^{II} ROA signal (1330 cm^{-1}) gives a strong, sharp positive ROA band, but the C_N-H^{II} (1355 cm^{-1}) vibration appears to give a sharp negative band. Significant contributions to the latter occur from the from the methyl symmetric deformation, which also gives negative ROA. (In Ala_2/NaOD the C_N-H^{II} deformation Raman band 'disappears' but the negative ROA signal remains.)

We now consider the NaOD and DCl spectra because it is more convenient to examine the decoupled spectra first. Upon raising the pD of the deuterated solvent the C_N-H deformations disappear, and it seems likely that these vibrations drop in frequency and become indistinguishable from the C_C-H deformations. Similar behaviour is found in alanine where deprotonation of the amino group lowers the C-H deformation frequency by 20 cm^{-1} (Gargaro, 1991).

In DCl we see similar behaviour to the of D_2O: the C_C-H deformations 'disappear' and it is difficult to pin down these new frequencies. The ROA studies on alanine propose that the two C-H deformations form a new band at 1330 cm^{-1} at low pH (Gargaro, 1991). This may be occuring here with the two new low frequency C_C-H deformations contributing to the weak C_N-H^I vibration. In DCl there is another band at 1255 cm^{-1} (which also occurs in alanine at low pH, but is assigned to O-H deformations). It is tempting to assume that the change in the C terminal residue's ionisation state will lower the C-H deformations (in a similar way as the N terminal ionization state is related to C_N-H deformation frequency) and assign this 1255 cm^{-1} band to the C_C-H^I deformation.

The Raman spectrum of Ala_2/HCl is essentially the same as that of Ala_2/H_2O (described above), except that the amide III^2 band vanishes and

another band appears at 1295 cm^{-1} (as a slight shoulder). This drop in frequency of the Raman band is accompanied by a similar change in the ROA, with the 1340 cm^{-1} band in H_2O splitting into an intense positive band at 1295 cm^{-1} and a smaller positive band at 1340 cm^{-1}. The deutereted solvent spectra show that the C_C-H deformations are sensitive the the carboxylate group ionization state, and it appears that the carboxylate protonation perturbs the amide III^2 vibration and lowers it. (It is possible that it is actually the amide III^3 that is lowered but, according to the normal mode analysis, it contains no C_C-H^I contributions.) The negative ROA peak assigned to C_C-H^I and amide III^1 drops by only ~10 cm^{-1} implying again that the C_C-H^I band is not as pure as originally thought.

When we go to NaOH the Raman and ROA spectra are not largely different from the H_2O spectra. Initially we would expect the amine deprotonation to drastically change the amide III^2 and amide III^3 vibration because these have large C_N-H^{II} contributions, but this is not the case in either the Raman or the ROA. The similarity between the ROA of Ala_2 in neutral and basic solvents, and the difference between the ROA of Ala_2 in neutral and acidic solvents, indicates that, although C-H deformation depends on terminal ionization, the amino terminal has less of an influence on the amide III vibrations than the carboxyl terminal. (The possibility of conformational changes are discussed in section 4.5.) The sharp 1340 cm^{-1} positive ROA band remains the same, but the 1270 cm^{-1} ROA negative peak changes to a couplet. Again this indicates that the ROA in the 1270 cm^{-1} region has more contributions than just the C_C-H^I vibration. The analysis of the Raman spectrum is hampered because the background in the amide III region of the NaOH sample is high.

In the hydrated solutions there is a small shoulder at ~1230cm^{-1}. This is a NH_3^+ rocking mode (Oboodi et al., 1984).

1000-1200 cm^{-1}

It is in this region that NH_3^+ rocks dominate the alanine vibrations, and the same is generally true for Ala_2. The acidic and neutral Raman and ROA spectra are almost identical. When we compare the neutral and alkali spectra we see that sharp changes occur in ROA for both hydrated and deuterated solvents and the Raman of hydrated solvents. However in deuterated solvents the Raman spectra remains similar because the ND_3^+ to NH_2 transition has less of a vibrational effect on this region (ND_3^+ rock occurs at 880 cm^{-1} according to Kettle et al., 1990).

Generally the assignments for this region become difficult. However some plausible explanations are certainly available.

In Ala_2 in H_2O and HCl there ia a negative/positive/negative feature between 1080 and 1200 cm^{-1}. This is made up of negative/positive and positive/negative couplets, and upon deuteration the low frequency couplet falls in frequency revealing the negative/positive/positive/negative feature seen in DCl and D_2O. The lower couplet changes upon deprotonation of the amino group and is assigned to CH_3 rock with NH_3^+ rock contributions. Ab intio calculations on alanine produce similar assignments in this region (Barron et al., 1991; Gargaro, 1991). The negative component of the high frequency (positive/negative) couplet in acidic and neutral solutions falls by ~15 cm^{-1} between the hydrated and deuterated solvents, while the low frequency component remains unchanged. This frequency difference is not present in the alkali solutions and decreases as we go up the series. This would appear to be an effect caused by the amino terminal, and this is confirmed by the changes that we see in the Raman

spectra between the neutral and alkali solutions. The vibration from the amino terminal that causes these effects may be NH_3^+ rock in hydrated solutions, and ND_3^+ deformations in deuterated solutions. In crystalline alanine the ND_3^+ deformation appears at ~1190 cm^{-1} (Kettle et al., 1990).

The Raman spectrum in this region is substantially different in NaOH than H_2O. Although we would expect the frequencies of the vibrations to be different, the situation is compounded by the convergence of the bands and appearance of a NH_2 twist at ~1080 cm^{-1} which occurs in alanine (Gargaro, 1991).

In HCl and H_2O the Raman band at 1000 cm^{-1} gives slightly negative ROA, in NaOH this band moves up to 1010 cm^{-1} and loses any ROA intensity. A similar band is also found in alanine: the zwitterionic form of alanine has a 1000 cm^{-1} Raman band with a large negative ROA and the high pH form has a ~1050 cm^{-1} Raman band with no ROA intensity. In alanine this vibration has been assigned to NH_3^+ rock with small CH_3 contributions (Barron et al., 1991 and 1992b). It seems reasonable to adopt this alanine assignment for Ala_2.

In DCl and D_2O the band at ~995 cm^{-1} can be assigned to N-D deformations, which are isolated (Krimm and Bandekar, 1986) and show no ROA. However, in NaOD the frequency falls by ~10cm^{-1} and acquires negative ROA.

1000-700 cm^{-1}

Between 900 and 960 cm^{-1} there are two peaks that give positive ROA in all six spectra. The Raman and ROA are identical in HCl and H_2O, but this is not true for DCl and D_2O where the 930 cm^{-1} Raman band is more intense at neutral pH. The ROA of these bands change most drastically

between neutral and alkali pH. The ROA of alanine shows a strong positive band in this region at all pH's. Although these bands have been assigned to CH_3 rock in alanine (Oboodi et al., 1984) and Ala_3 (Qian et al., 1991), the ROA data of alanine and other amino acids support C_α-C(=O) stretch vibrations.

There are two Raman bands between 800 and 900 cm^{-1}. They are both lowered by ~5 cm^{-1} upon deuteration and the weaker, low frequency vibration, which gives a negative ROA signal, is pH dependent. A strong Raman band, with negative ROA, and a similar frequency dependence upon pH as the weak, low frequency vibration is seen in alanine where the band is assigned to CCN stretch with a CO_2^- symmetric bend component (Kettle et al., 1990; Oboodi et al., 1984; Gargaro, 1991). Qain et al. (1991) assign a band similar to the more intense, high frequency band in this region to CCN stretch across the amide link. This vibration certainly originates in the amide link because there is no equivalent band in alanine and there is no pH dependence. The issue is complicated in deuterated solvents by the presence of ND_3^+ rocking components which occur about 880 cm^{-1} (Kettle et al., 1990).

A Raman band at ~740cm^{-1} under acidic conditions, moves up to 770 cm^{-1} in neutral and alkali conditions (although in D_2O a smaller peak remains at 740 cm^{-1}). This feature consistently gives a negative ROA and draws parallels with the alanine ROA where a similar pH/frequency dependence is observed. Therefore we assign the 770 cm^{-1} band to CO_2^- wag (Gargaro et al., 1993) and possibly CO_2^- symmetric bend (Barron et al., 1991) which is found in all the zwitterionic amino acids. The 740 cm^{-1} band is predicted by ab initio calculations on neutral alanine as arising from C=O out-of-plane bend and O-H torsion (Barron et al., 1991).

700-350 cm^{-1}

The low frequency vibrations are difficult to assign because the vibrational coordinate contibutions to the normal modes are very structure specific. Normal mode calculations on Ala_3 (Qian et al., 1991) show very different low frequency normal modes depending on the conformation adopted by the Ala_3. However, it can be seen that the low frequency ROA changes only slightly upon deuteration. The strong negative ROA band may originate from C=O in plane bends which also give negative ROA in alanine dipeptide (see chapter 5).

3.3.3 Ala_3

The Raman and ROA spectra of Ala_3 in hydrated and deuterated solvents are shown in Figures 3.7 and 3.8.

1750-1400 cm^{-1}

The Raman spectrum of Ala_3 is similar to that of Ala_2, but the analysis in the hydrated species is made more difficult by the increased relative intensity of the H_2O bend Raman band at 1650 cm^{-1}. The ROA spectra are also similar.

In Ala_3/H_2O the amide I vibration gives a Raman band at 1680 cm^{-1} with a positive ROA signal (similar to Ala_2). Another amide I vibration may be present at 1650 cm^{-1} but this could be the broad H_2O band. Two amide I' bands can be seen in D_2O, one being a low frequency shoulder of the other, both bands carry positive ROA. In H_2O the Raman band appearing as a shoulder at 1575 cm^{-1} is probably the carboxylate asymmetric stretch or the amide II. As we go to alkaline pH we see that in NaOH the previously distinct amide I band has vanished under the H_2O band, again

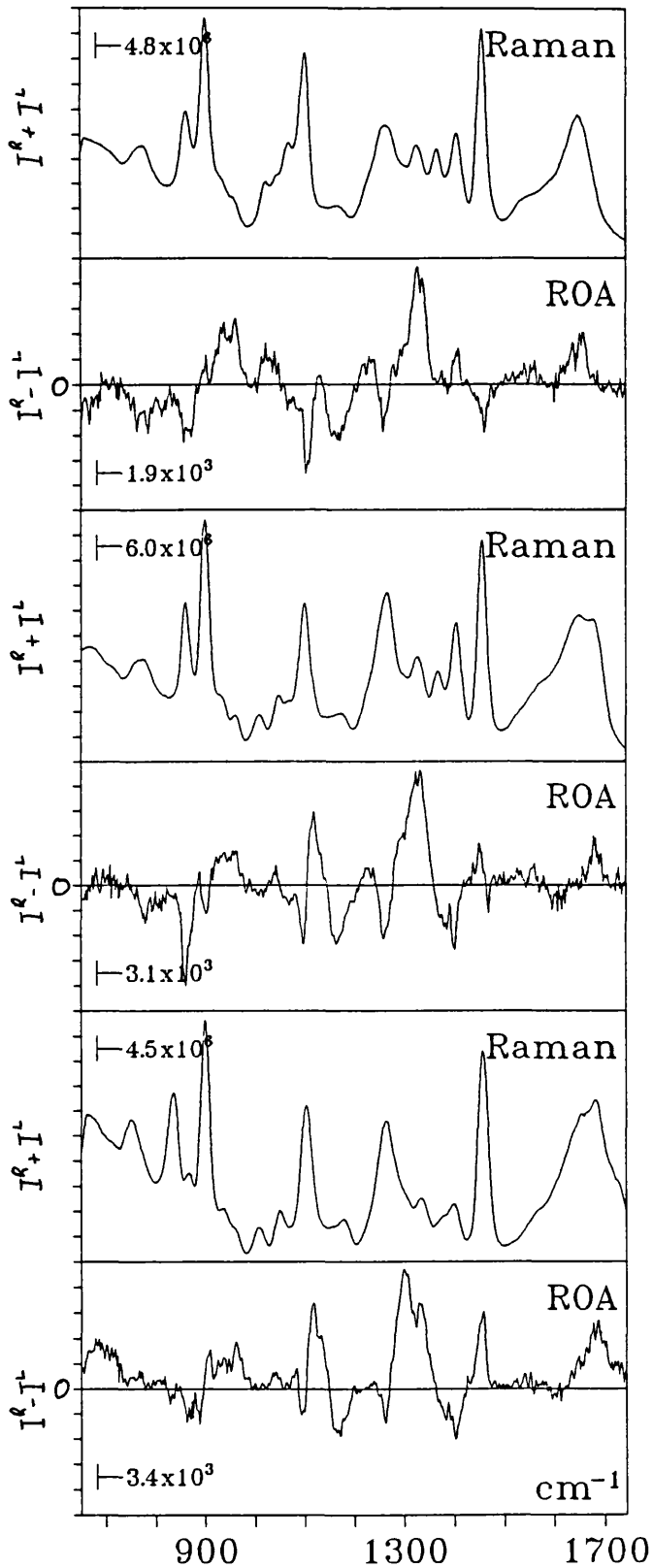


Figure 3.7: Ala_3 in hydrated solvents: 1M NaOH (top), H_2O (middle), 1M HCl (bottom).

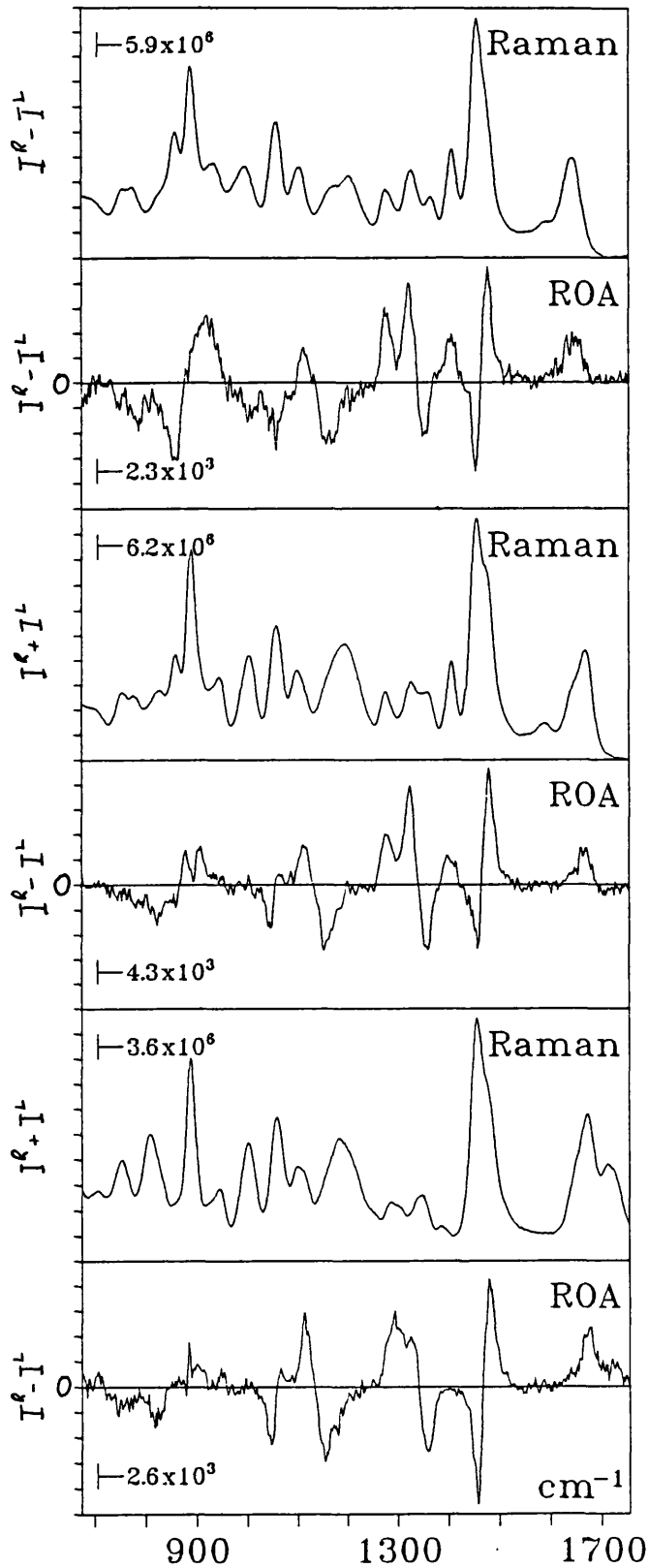


Figure 3.8: Ala_3 in deuterated solvents: 1M NaOD (top), D_2O (middle), 1M DCl (bottom).

the amide I band gives positive ROA. As with all the Ala_3 in hydrated solvents a weak ROA signal is seen between 1510 and 1570 cm^{-1}; this corresponds to the amide II band. It is a possiblity that the positive ROA could be due to the COO^- asymmetric stretch, however, this argument is weakened by the fact that it occurs in Ala_3 in HCl, and is absent in Ala_3 in NaOD and D_2O.

In DCl the amide I' bands appear as a strong peak at 1675 cm^{-1} and a shoulder at 1650 cm^{-1}, and as with D_2O both these bands give positive ROA. In both the acidic solvents the carboxylic acid C=O stretch appears at 1700 cm^{-1} and give positive ROA. (Although these Ala_3 C=O stretch Raman bands are smaller than in Ala_2, the ROA is more pronounced. This is probably an artifact or an effect of the intensifying amide I(I') bands, because further up the series the Ala_4 and Ala_5 species do not show any ROA on the COOH(D) band at all.) The NH_3^+ assymmetric deformation is too weak to be seen in the Raman of Ala_3 in HCl or H_2O, but may be responsible for the slightly negative ROA peaks at 1610 cm^{-1}.

Our results brought us into conflict with previous VCD studies on Ala_3 that reported a conformational change between Ala_3 at neutral pD and Ala_3 at alkaline pD (Lee et al., 1989). Therefore we carried out a more detailed depolarization ratio study of the amide I region of Ala_2 and Ala_3 in D_2O and NaOD.

VCD and IR spectra obtained by Lee and co-workers of the amide I' of Ala_3, at pH=7, show a doublet and a distinct negative/positive VCD couplet at pH=7, while at high pH only a single IR peak with no VCD is observed. It was proposed that the VCD couplet was caused by the coupling of the two individual amide I' vibrations in accordance with the coupled oscillator model (Lee et al., 1989). Using theoretical calculations a series of conformations that would produce such a VCD couplet were

established and molecular mechanics calculations were used on these conformations to propose an Ala_3 structure. Lee and co-workers proposed that in the zwitterionic form the electrostatic interactions of the endgroups conferred a stable, 'biting tail' structure on Ala_3. When the pH was raised, the charge altered, and the molecule was 'released' producing a conformational fluctuating molecule which would not sustain a coupled oscillator mechanism.

There were several features in our Raman and ROA spectra that caused us to question this hypothesis: the amide III ROA did not indicate large conformational changes between neutral and alkaline pH; the Raman doublet that appears in the amide I' region of zwitterionic Ala_3 also appears in the DCl spectrum; and the Ala_2 amide I' Raman band frequency is 25 cm^{-1} lower in NaOD than in D_2O easily explaining any couplet collapse (this can also be seen in the Ala_2 H_2O/NaOH spectra).

One of the features of the 'coupled oscillator' system that was readily amenable to study was the Raman depolarization ratios (when two vibrating groups coupling in this way they form a symmetric and antisymmetric pair which have ideal depolarization ratios of $0 \leq \rho \leq 0.75$ and 0.75 respectively.) New samples of Ala_2 and Ala_3 in D_2O and NaOD were made up to the original concentrations and the resultant polarized and depolarized spectra were deconvoluted (using the Curvefit analysis in the Lab Calc software package).

During the deconvolution of both Ala_2 and Ala_3 at neutral pH, two amide I' bands were found. The lower frequency amide I' band of Ala_2 had a similar frequency and depolarisation ratio to the 1645 cm^{-1} band in Ala_3 at pH 7 and probably represents a low population conformer. The Raman depolarisation ratios of vibrations between 1550 and 1780 cm^{-1} are presented in Table 3.3.

Table 3.3: Raman depolarization ratios of Ala_2 and Ala_3 in D_2O and NaOD.

		CO_2^- asymmetric stretch (1585 cm^{-1})	Amide I' (low frequency) (1645 cm^{-1})	Amide I' (high frequency) (1665 cm^{-1})
Ala_2	D_2O	0.74	0.20	0.13
	NaOD	0.66	0.17	-
Ala_3	D_2O	0.73	0.20	0.10
	NaOD	0.58	0.14	-

These results indicate that the origin of the IR, Raman and VCD signals in the zwitterionic Ala_3 is simply the amide I' modes from different amide links. The zwitterionic Ala_3 depolarization ratios do not sustain a coupled oscillator model. The singlets, and the collapsed VCD, in the high pD Ala_3 are caused by the superposition of the high frequency amide I' vibration on the lower one (this frequency change occurs in Ala_2). These data alone do not preclude a conformational change (the population of conformational substates that are occupied probably does alter) but the ROA data appear to indicate that the large conformational changes proposed by Lee and co-workers do not occur. It was noted by Lee et al. (1989) that the absence of solvent in the original molecular mechanics calculations could induce a false conformation and we believe that this is the case. Experimental NMR evidence also indicates that the zwitterionic interaction has an insignificant effect on conformation due to aqueous solvation (Sekacis et al., 1988).

The asymmetric methyl deformations occur in all six spectra at ~1460 cm^{-1}, and have similar Raman and ROA characteristics to the methyl deformations in Ala_2. The only exception is the positive asymmetric methyl deformation ROA in the Ala_3/HCl sample which was a small couplet in Ala_2/HCl. It is not obvious why this has changed and one plausible explanation is a positive artifact, on what is quite a strongly polarized

band. The ROA spectra of Ala_4 and Ala_5 in HCl show similar features to the Ala_3.

The 1410 cm^{-1} Raman band in the alkali and neutral solutions is the carboxylate symmetric stretch. However, in the acidic solutions bands still appear in this region due to methyl symmetric bends (Qian et al., 1991). The disappearance of the 1405 cm^{-1} band between Ala_3/HCl and Ala_3/DCl is due to a N-H in plane bend (Qian et al., 1991).

In the deuterated solvents the ROA assignments are straightforward. In DCl there is no ROA or Raman band at 1410 cm^{-1}; in D_2O and NaOD the new COO^- Raman band carries with it positive ROA. (This is similar to the situation in Ala_2 for both deuterated and hydrated solvents.) For Ala_3 and Ala_4 in the hydrated solvents, this region appears to be more complicated. It is unclear whether these complications are ROA artifacts or some methyl symmetric bend and N-H in plane bending effect which predominates over COO^- vibrations in higher members of the series.

1400-1200 cm^{-1}

The Raman and ROA spectra of the 1420-1200 cm^{-1} region are shown in Figure 3.5 and 3.6.

We are going to change the nomenclature of the amide bands so as not to confuse the Ala_3 results with the Ala_2. The Raman and ROA spectra can be explained nicely by extrapolating from Ala_2 and remaining with three amide III vibrations. So we call the vibrations amide III^a, amide III^b and amide III^c (these are related to the amide $III^{1,2,3}$ of Ala_2 only in the spectral, or experimental, properties of the bands).

As with the other peptide spectra we turn to the uncoupled, deuterated results first. The Raman and ROA features of Ala_3 in the deuterated

solvents are very similar to the analogous Ala_2 samples. This contrasts with the equivalent results in the hydrated solvents where the Raman and ROA features are different, but is not unexpected as the N-D in plane bend does not allow the mixing of the C-H deformations across the amide link. It also shows that the decoupling of the amide III vibration simply produces a summation of the individual C-H vibrations.

The Raman and ROA features remain the same between Ala_2/D_2O and Ala_3/D_2O and so we take our assignments from the Ala_2 spectra. The 1280 cm^{-1} Raman band, which will be a combination of various $C-H^I$ deformations becomes slightly more intense and broad in Ala_3 than in Ala_2. The ROA becomes more intense reflecting the additional $C-H^I$ deformations. The 1325 cm^{-1} Raman band increases in intensity with respect to the Ala_2/D_2O, and again this indicates additional $C-H^{II}$ deformations. The 1355 cm^{-1} Raman band which was assigned to C_N-H^I in Ala_2 decreases, reflecting the decreasing influence of the end groups. The positive/negative couplet at 1325 cm^{-1} to 1360 cm^{-1} can be assigned to $C-H^{II}$ bends, with the higher frequency component having a significant contribution from the methyl symmetric deformation.

The Raman spectra of Ala_2 in DCl and Ala_3 in DCl are very similar; slight broadening occurs on the 1345 cm^{-1} band, and another peak appears at 1280 cm^{-1} in Ala_3, probably due to additional C-H deformations. The 1280 cm^{-1} Ala_3/D_2O band moves up to 1290 cm^{-1} in DCl and another band appears at 1300 cm^{-1}; both these bands carry positive ROA and so are probably related to the $C-H^I$ deformations. As with Ala_2 the 1325 cm^{-1} Raman band, and its positive ROA, collapse in DCl leaving some small (residual) positive ROA. This collapse is also seen in the Ala_2 spectra, where the band concerned was assigned to the C_C-H^{II} vibrations and we adopt this assignment for Ala_3.

When we compare the Ala_3/D_2O spectrum to the Ala_3/NaOD spectrum we see that the changes in the ROA spectra are similar, but less pronounced, than the comparison between Ala_2/D_2O to Ala_2/NaOD. The same collapse of the 1365 cm^{-1} Raman band, assigned to the N-terminal $C-H^{II}$ bends, the drop in positive ROA intensity at 1305 cm^{-1}, assigned to C_N-H^I, with the subsequent increase in positive ROA at 1280 cm^{-1} were all seen in Ala_2 spectra.

If we compare Ala_2/H_2O to Ala_3/H_2O then we see that in the Raman spectrum the broad 1275 cm^{-1} combination band drops in frequency by 10 cm^{-1} and an additional shoulder appears at about 1255 cm^{-1}. The former is the amide III^a band which is lower than the amide III^1 band due to increased delocalization, and the latter is probably additional C-H deformations caused by the extra residue. The amide III^2 and amide III^3 features in Ala_2 are replaced by a single Raman peak at 1329 cm^{-1} in Ala_3. The sharp positive ROA peak at 1340 cm^{-1} in Ala_2 is lower in frequency in Ala_3. Again this is probably caused by the amide III^b and amide III^c bands which fall in frequency due to the increased delocalization.

The Raman spectra between 1200 and 1400 cm^{-1} of Ala_3 in H_2O and HCl remains similar, with the exception of the CO_2^- and CH_3 bands. When we turn to the ROA spectra we see a substantial change between Ala_3/H_2O and Ala_3/HCl. Again, similar to the situation in Ala_2, the change in endgroup charge appears to have split apart the amide III^b and amide III^c bands, taking ROA intensity from the 1335 cm^{-1} band of Ala_3/H_2O and depositing it at ~1300 cm^{-1}. The drop in amide III^b ROA is 10 cm^{-1} less than the drop in the Ala_2 amide III^2 because; firstly, the COO^- to COOH transition will have less of an effect on the vibrations of the molecule, and secondly, the more delocalized vibrations will quench the effect.

The Raman and ROA spectra stay almost identical between H_2O and NaOH with the exception of the 1265 cm^{-1} Raman band, which drops in intensity. This is seen in Ala_2 as well, and reflects the redistribution of the $C-H^I$ vibrations at the amino terminal.

1200-1000 cm^{-1}

The differences between the Raman spectra of the di- and tripeptide in the individual solvents in the rest of the region are still quite small. In the aqueous solvents the 1110 cm^{-1} Raman band increases in the trimer; this concurs well with the Ala_2 assignment of CH_3 rock and CH bends. Again this band splits in two in deuterated solutions, which would indicate that in hydrated solvents the bands had a $N-H/NH_3^+$ component. The Raman bands do not change when we go from neutral to alkali in either solvent, indicating that an N-H contribution is applicable rather than an NH_3^+ contribution. As with the Ala_2, Ala_3 in hydrated solutions shows a Raman band at 1075 cm^{-1} in alkali solutions. This maybe NH_2 twist, which as we would expect, is not as intense in Ala_2.

If we examine the ROA of the region we again see a negative/positive/negative feature between 1050 and 1200 cm^{-1} in the HCl and H_2O solutions. However in deuterated solutions the feature now becomes negative/zero/positive/negative (i.e. the low frequency couplet has lost it's positive contribution). The lower 'couplet' is obviously still deuterium sensitive, but in a different way to that found in Ala_2. Therefore we cannot carry the Ala_2 assignment straight over to Ala_3 and the assignment of the low frequency couplet in Ala_3 remains unclear. The highest frequency negative component remains constant between all the hydrated and deuterated solvents is assigned to CH_3 rocks and C-H bends (there is no NH_3^+ contribution because there is no deuterium dependence).

In Ser_3 we do not see this feature at all indicating that the CH_3 rock is the dominant contribution.

A Raman band occurs at 1000 cm^{-1} in H_2O and HCl and has no ROA. An NH_3^+ rock dependence is demonstrated by the ~15 cm^{-1} shift to higher frequency that occurs at high pH, and it seems reasonable to assume that the band assignment is similar to the corresponding band in Ala_2; CH_3 rock vibration with a slight NH_3^+ rock contribution. This assignment agrees with Qian and co-workers (1991), but they propose that the vibration is dominated by C-C stretch. The ROA characteristics of this band in Ala_3 are slightly different than in Ala_2 or Ala_4.

In the D_2O and DCl the 1000cm^{-1} band is assigned to N-D in plane deformations, the band carries no significant ROA. As with Ala_2 the band drops slightly at high pD, and acquires a negative ROA peak.

960-750 cm^{-1}

In hydrated solvents there are two bands between 900 and 960 cm^{-1}, their frequency and intensity remain quite similar over the pH scale. Similarly, in deuterated solvents the single band at ~950 cm^{-1} remains constant.

Normal mode calculations suggest that the bands at around 920 to 940 cm^{-1} are due to C_α-N stretch contributions, while higher up at ~960 cm^{-1} methyl rocks prominate. Qain and co-workers (1991) note that several N-D in plane deformations (the so called amide III' band) will occur between 900 and 1000 cm^{-1} and these may account for the ROA sharpening up in the deuterated samples.

Between 800 and 900 cm^{-1} a series of Raman bands appear that show similar characteristics to the bands found in Ala_2. Two bands appear in

H_2O at 850 and 900 cm^{-1} but when the carboxylate group is either protonated or deuterated the low frequency band drops by ~20 cm^{-1}. In HCl, DCl, D_2O and NaOH the low frequency band is composed of two bands. As in Ala_2 we assign the low frequency vibration to CCN stretch with COOH contributions. The situation is complicated in deuterated solvents by ND_3^+ rock contibutions which occur about ~880 cm^{-1}. The high frequency band remains virtually constant in all six spectra, but has a slight deuterium dependence. The high frequency vibration is C-N stretch across the amide link. Qain and co-workers (1991) show that this band comes from C-N stretch and does have some deuterium dependence.

The single 750 cm^{-1} vibration has very similar Raman and ROA pH dependence to the equivalent band in Ala_2. This would suggest a similar assignment of COO^- wag and bend. The small magnitude of the ROA is consistent with the idea that the 770 cm^{-1} band belong to an end group. It has also been suggested that ND_3^+ contributions may play a part in the lower frequency band (Bandekar and Krimm, 1988).

3.3.3 Ala_4 and Ala_5

The analysis of Ala_4 and Ala_5 are combined, mainly because the quality of the spectra is not good enough to carry out a thorough analysis. The poor quality is due to the low solubility of these longer alanyl peptides. The Raman and ROA spectra of Ala_4 and Ala_5 are shown in Figures 3.9 to 3.11.

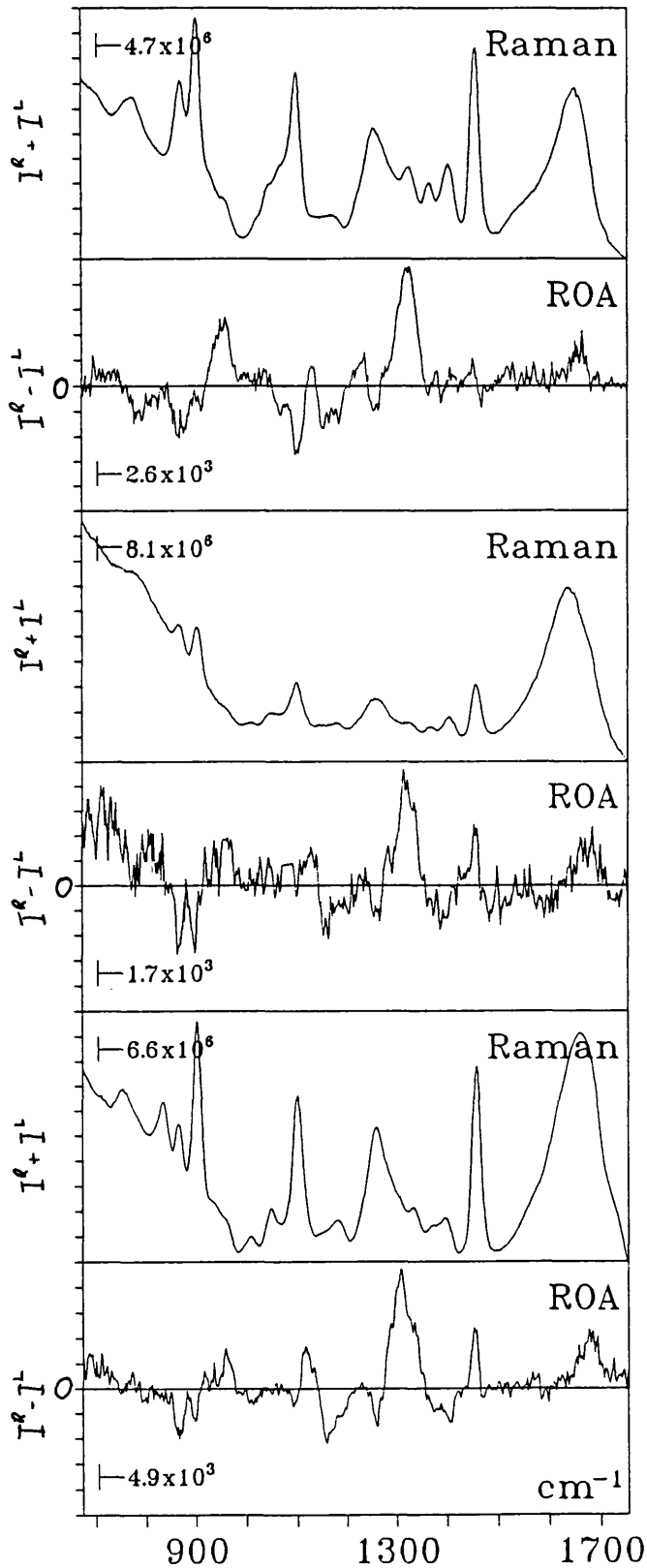


Figure 3.9: Ala_4 in hydrated solvents: 1M NaOH (top), H_2O (middle), HCl (bottom).

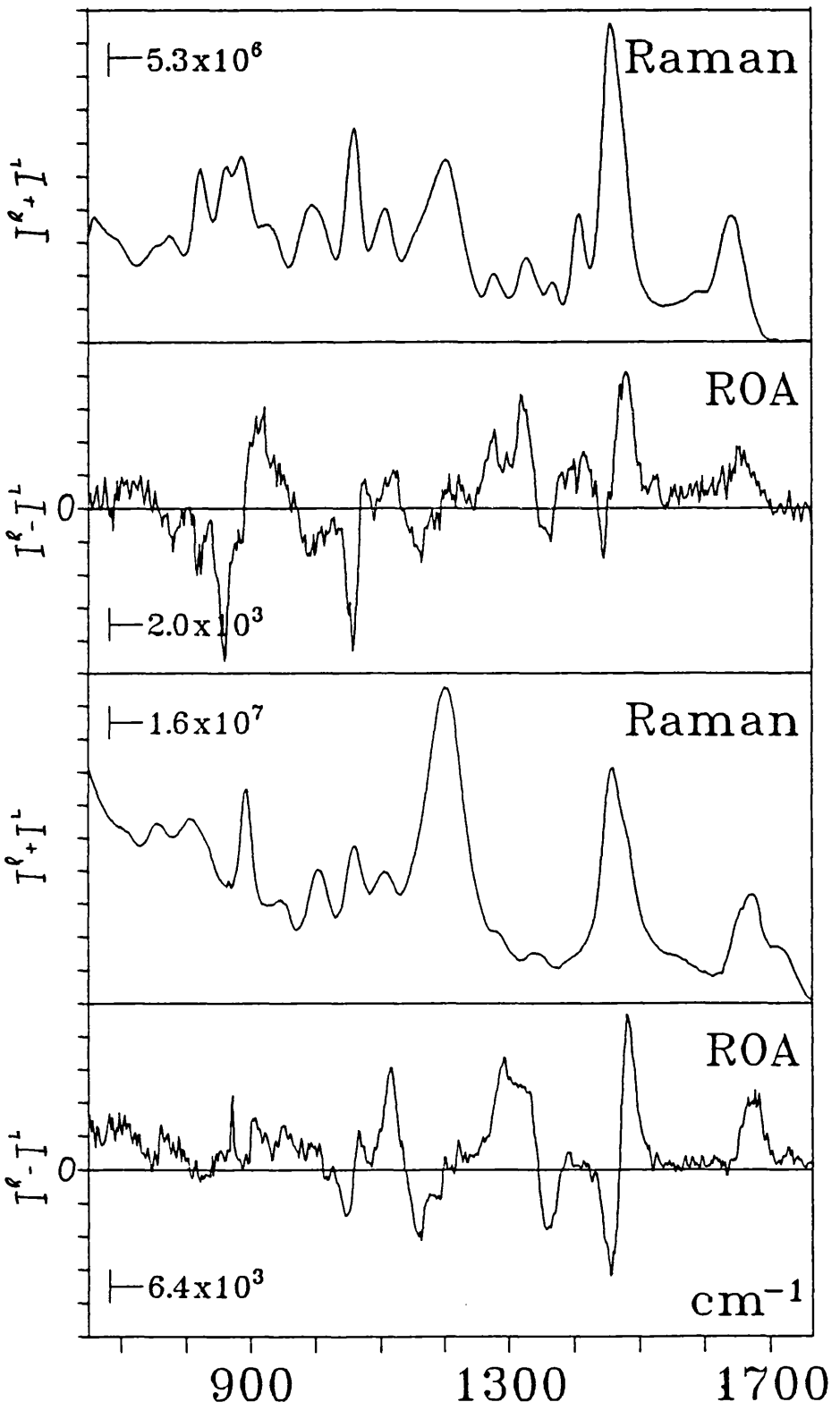


Figure 3.10: Ala_4 in deuterated solvents: 1M NaOD (top), 0.2M DCl (bottom)

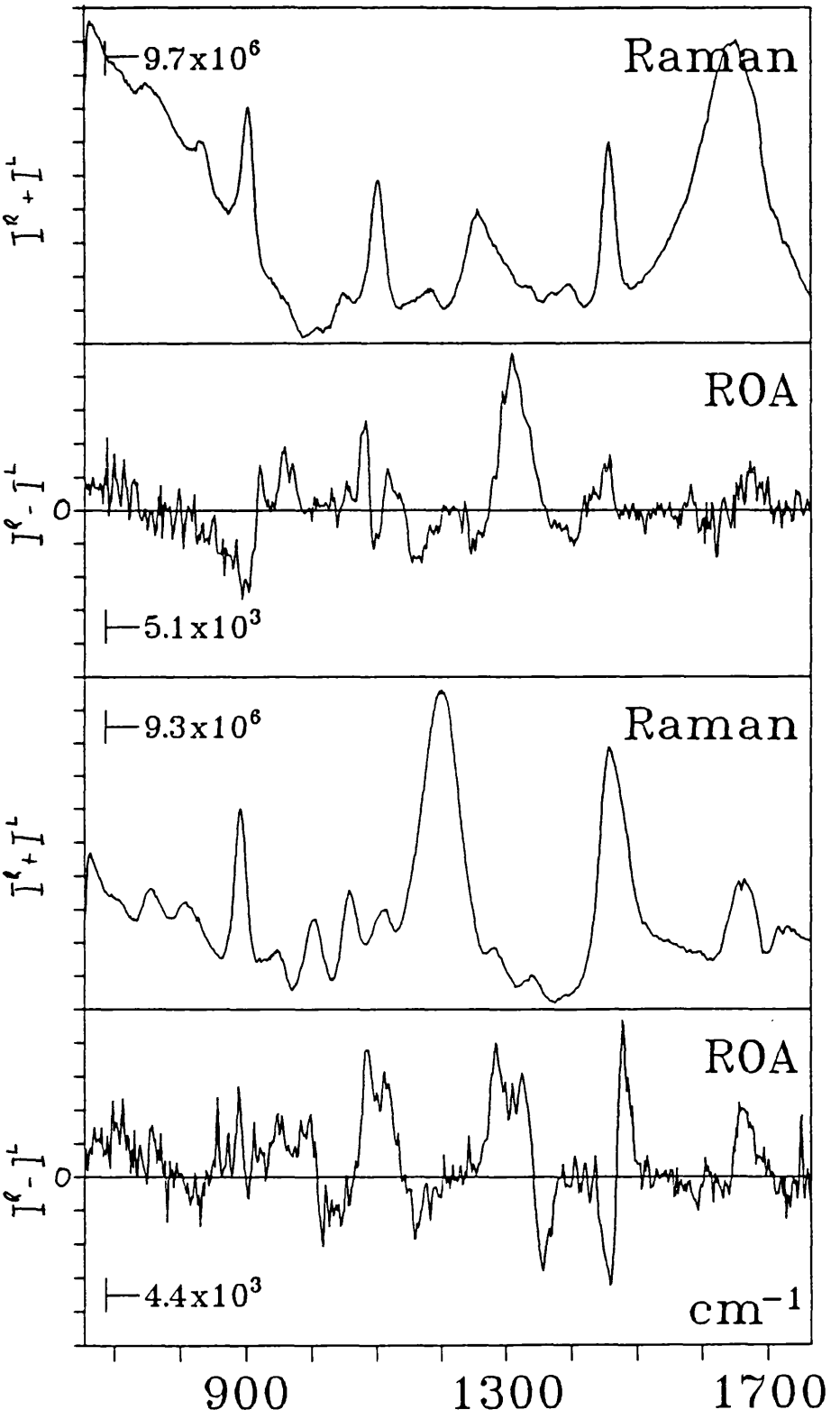


Figure 3.11: Ala_5 in 0.1M HCl (top), 0.1M DCl (bottom)

1750-1400 cm^{-1}

The amide I Raman vibration of Ala_4 in hydrated solvents are almost completely swamped by the Raman band of the H_2O bend. In HCl and NaOH the amide II and/or carboxylate asymmetric stretch vibrations can be seen as weak low frequency shoulders about 1570 cm^{-1}. The situation is far clearer in the deuterated solvents, where the 1720 cm^{-1} carboxylic acid C=O stretch, the amide I' band (at 1675 cm^{-1} in DCl and 1640 cm^{-1} in NaOD) and the 1580 cm^{-1} COO^- asymmetric stretch are all discernible. The amide I and I' ROA signals are all positive, but the frequency is 20 cm^{-1} lower in the alkali solvent than in acidic. The Ala_4 and Ala_5 spectra are very similar.

The asymmetric methyl deformations and amide II' vibrations appear at 1460 cm^{-1} in Ala_4 and Ala_5. The Raman and ROA features are the same as those in the Ala_3 spectra and so are not discussed any further.

1400-1200 cm^{-1}

The Raman and ROA features of the carboxylate stretch and methyl symmetric deformations in Ala_4 and Ala_5 are the same as those in the equivalent Ala_3 spectra. The explanations of the various feature are the same as with the Ala_3 spectra. The carboxylate stretch ROA in Ala_4/NaOD is smaller than that in Ala_3 due to the diminishing effect of the end groups.

In the amide III region the Ala_4/H_2O Raman and ROA spectra have the same overall features as the Ala_3/H_2O ROA. The 1410 cm^{-1} carboxylate symmetric stretch still remains, and gives negative ROA, but both the Raman and ROA intensities are smaller because the endgroup will have a diminishing effect. (The 1305 cm^{-1} Raman band which appeared started to

appear in Ala_3 in hydrated solvents appears again, and may be more intense, in Ala_4 in hydrated solvents.) The amide $III^1/C_\alpha\text{-}H^I$ deformation-equivalent Raman band drops in frequency by 5 cm^{-1} between Ala_3 in hydrated solvents and Ala_4 in hydrated solvents. The positive amide $III^{2/3}$-equivalent ROA signal has dropped in frequency by 10 cm^{-1} between Ala_3 and Ala_4 in H_2O and NaOH. This is caused by the more extensive delocalization of the amide III vibrations along the longer peptide backbone.

The ROA of Ala_2 in HCl gave two positive peaks at 1295 and 1340 cm^{-1}; Ala_3/HCl showed two broad positive ROA peaks at 1300 and 1335 cm^{-1} with the latter of the latter being smaller; the trend continues in Ala_4/HCl where the high frequency band vanishes and only a single positive signal appears at 1310 cm^{-1}. The shift in the amide $III^{1/b}$-equivalent between the alanyl peptides in H_2O and HCl gradually decreases as the series increases: 45, 35 and 10 cm^{-1} in Ala_2, Ala_3 and Ala_4 respectively. The Ala_4/HCl and Ala_5/HCl Raman and ROA spectra are almost identical.

The Ala_3/DCl, Ala_4/DCl and Ala_5/DCl ROA spectra are almost identical. However, the similarity between Raman spectra is not so obvious because the concentration of Ala_4 in DCl is so low. The Raman and ROA spectra of Ala_3/NaOD and Ala_4/NaOD are identical.

In the deuterated samples the strong band at about 1200 cm^{-1} is the D_2O bend.

1200-1000 cm^{-1}

The Raman and ROA spectra of Ala_4 in hydrated solvents are similar to their Ala_3 equivalents. The 1000 cm^{-1} Raman band appears to decrease in intensity relative to the surrounding Raman bands, probably due to the NH_3^+ dependence which has been noted previously.

The Raman and ROA spectra of Ala_3, Ala_4 and Ala_5 in deuterated solvent are all very similar except for the 1100 cm^{-1} ROA band which has been distorted by artifacts in the DCl and the 1050 cm^{-1} artifact in the ROA spectrum of Ala_4/NaOD.

1000-750 cm^{-1}

The Raman and ROA spectra of Ala_4 and Ala_5 in HCl are almost identical to the Raman spectrum of Ala_3 in HCl. A difference occurs in the Raman band at 850 cm^{-1} which appears to increase in intensity between Ala_3 and Ala_4, but vanishes in Ala_5. The 950 cm^{-1} band in the Ala_5 spectrum is distorted by artifacts. The Raman of Ala_4 in H_2O is almost identical to that of Ala_3/H_2O. The ROA of Ala_3 and Ala_4 in H_2O are as similar as can be expected considering the poor signal-to-noise ratio of the latter. The Raman and ROA spectra of Ala_4 and Ala_3 in NaOH are very similar to each other.

The Raman and ROA spectra of Ala_3, Ala_4 and Ala_5 in DCl are very similar except for the relative intensities of the 750 and 800 cm^{-1} bands where the latter gets smaller as we go up the series. This concurs with our previous discussion of these bands which assigns the 750 cm^{-1} band to amide link C=O out of plane bends and wags and the 800 cm^{-1} band to CCN stretch and COOD vibrations. The ROA couplet in this region that

appears in Ala_2 and Ala_3 appears to be alot smaller in Ala_4 and Ala_5 (although the Ala_5 spectrum is poor quality).

The ROA spectrum of Ala_4 in NaOD is very similar to the spectrum of Ala_3 in NaOD, except for the artifact at 850 cm^{-1}. On the other hand the Raman spectra comparison shows many differences. The Ala_2, Ala_3 and Ala_4 spectra are all quite different, with six Raman bands appearing at 925, 885, 865, 825, 775 and 750 cm^{-1} in the latter.

3.4 CONCLUSION

By systematically studying a simple peptide series we have been able to:

- observe the delocalization of the amide III ROA along the peptide backbone and the dependence of this delocalization on the N-H in plane bend.

- observe and describe the changes to the C_α-H deformations and amide I vibrations with pH.

- assign other features of the alanyl peptide Raman and ROA spectra between ~1000-1800 cm^{-1}.

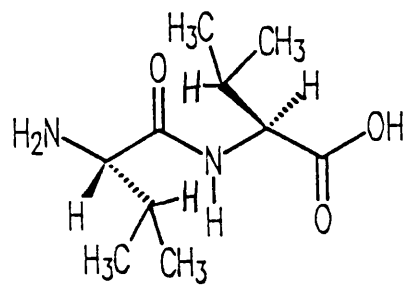
Chapter 4

Val_2 and Ser_3

4.1 INTRODUCTION

Once the alanyl oligopeptide series had been examined we turned our attention towards two other peptides, namely divaline (Val_2) and triserine (Ser_3). These are readily available, suitably soluble, and, although very few vibrational studies have been done on these oligopeptides, we can analyse their Raman and ROA spectra by using our previous ROA data from valine, serine and the alanyl peptide molecules (Gargaro et al., 1992; Ford et al., 1994). Naturally, different sidechains will change the vibrational spectrum of a solvated peptide, but it can do so in two different ways: first, the sidechains themselves could contribute different vibrational coordinates to the normal modes, and second, the sidechain may alter the conformational and dynamic behaviour of the molecule in solution. Seperating these two contributions to various vibrational spectra can be difficult and complicated. The structures of Val_2 and Ser_3 are shown in Figure 4.1.

(a)



(b)

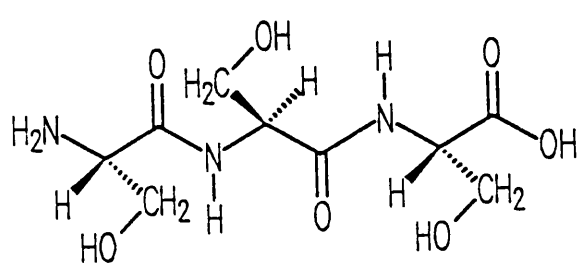


Figure 4.1: Structural formulae of Val_2 (a), and Ser_3 (b)

4.2 EXPERIMENTAL

The Val_2 and Ser_3 samples were bought from Sigma Chemical Co. and used without further purification. The same preparation protocol was used for the peptide samples as for the protein samples (see section 6.2.1).

Val_2 is soluble enough to be dissolved in H_2O and D_2O without having to raise, or lower, the pH. The deuterated Val_2 was prepared by lyophilization of a Val_2/D_2O solution. The concentrations for the H_2O and D_2O samples were both ~0.5M.

Ser_3 is not soluble at neutral pH, but a suitably high concentration can be obtained by dissolving it in HCl or NaOH. The Ser_3 ROA sample concentration was ~0.5M in 1M HCl. Deuterium exchanged Ser_3 proved difficult to obtain due to discolouration during the lyophilization process. Consequently, the Ser_3/DCl solution was prepared by dissolving Ser_3 straight from the bottle into the 1M DCl. (Only ~3% of the Ser_3 molecules will have more than 1 unexchanged proton.)

4.3 ASSIGNMENTS OF THE DIVALINE RAMAN AND ROA SPECTRA.

The Raman and ROA spectra of Val_2 in H_2O and D_2O are shown in Figure 4.2. Expanded amide III and amide III' Raman and ROA spectra are shown in Figure 4.3.

4.3.1 Overview

The Val_2 spectra generally appear to be better defined than the Ala_2 spectra and this may be related to the different properties of the sidechains. The larger and more hydrophobic valine sidechains may cluster together, inducing a more rigid solvated structure than that found in Ala_2:

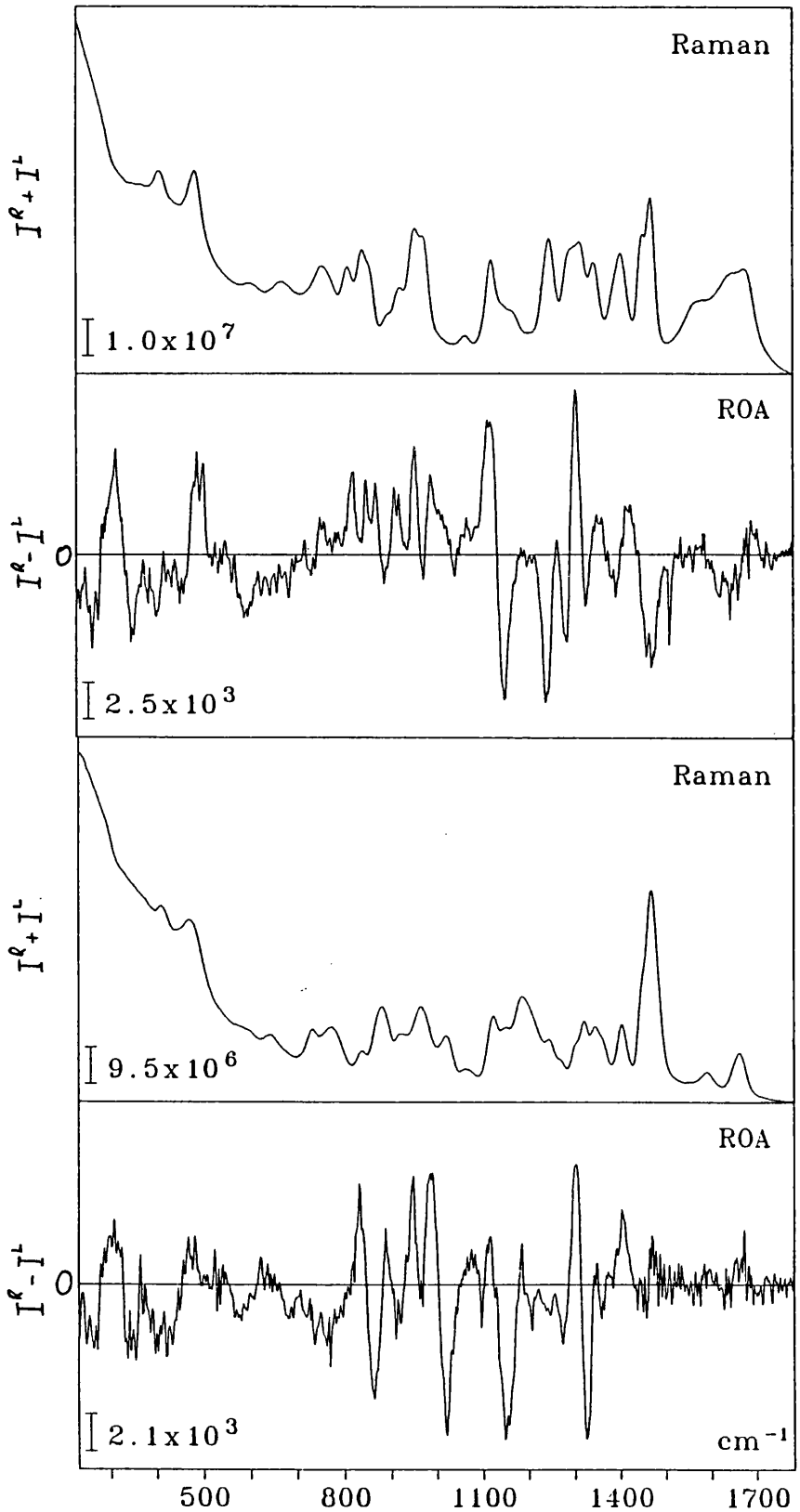


Figure 4.2: The Raman and ROA spectra of Val_2 in H_2O (top), and D_2O (bottom).

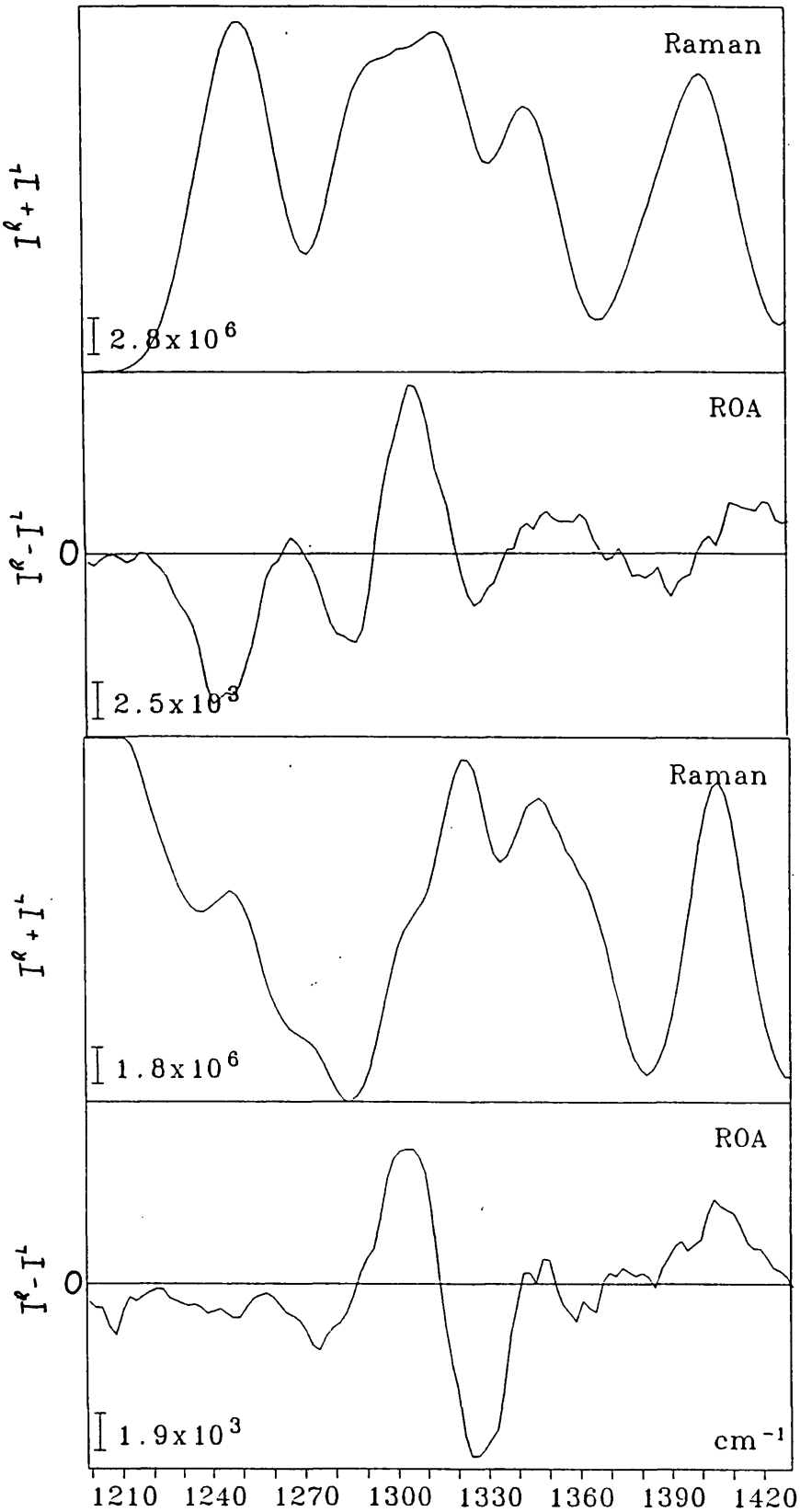


Figure 4.3: The Raman and ROA spectra of the 1200-1430 cm^{-1} region of Val_2 in H_2O (top) and D_2O (bottom).

the range of conformations occupied by the molecules in solution is probably smaller for Val_2 than Ala_2.

1700-1500 cm^{-1}

In Val_2/H_2O the amide I band occurs at ~1670 cm^{-1}, 10 cm^{-1} higher than in solid Val_2 (Fasman et al., 1978). In D_2O this band falls to 1665 cm^{-1}. We assign the 1630 cm^{-1} vibration in the H_2O sample to NH_3^+ deformation (Fasman et al., 1978). In Ala_2 this band occurs at 1625 cm^{-1}. At 1565 cm^{-1} there is a band that remains in D_2O, a similar band occurs in the Ala_2 samples (at 1570cm^{-1}) and we assigned it to COO^- asymmetric stretch (Fasman et al., 1978). The Val_2 ROA of this region is similar to that of the alanyl peptides; i.e. a couplet in H_2O and a doublet in D_2O (the negative NH_3^+ deformation ROA signal vanishes in D_2O).

The two peaks at 1440 cm^{-1} and 1460 cm^{-1} in H_2O can readily be assigned to the asymmetric CH_3 bends (Simons et al., 1972 and Gargaro et al., 1992). The Raman signal from the asymmetric methyl bends of valine are of equal intensity in both the solid and aqueous states (Gargaro, 1991 and Fasman et al., 1978), but in Val_2 spectra the high frequency peak has the greater intensity. The methyl vibrations give very strong ROA (stronger than the equivalent vibrations in Ala_2), possibly because there are twice as many methyl groups in valine as alanine, or because the molecule is being held in a more rigid conformation.

As with the Ala_2, the amide II' vibration couples with the methyl vibrations in D_2O and the resultant Raman band has a far greater intensity. (The increase is only slightly more going from Val_2/H_2O to Val_2/D_2O than it is going from Ala_2/H_2O to Ala_2/D_2O.) The interesting feature is that, in contrast to the alanyl peptides and Ser_3, the methyl ROA vanishes when the amide II' band couples. This may be because, in

valine alone, the vibrating groups are separated from the amide link (and the chiral backbone) by an extra atom. From this premise one could start to postulate various mechanisms such as the amide II' altering the normal modes in that region so as to remove the chiral component of the methyl vibrations, or a cancellation of two equivalent but opposite ROA signals from two separate, uncoupled vibrating groups (although the latter is unlikely).

The Raman band around 1400 cm^{-1} is readily assigned to the COO^- symmetric stretch. However, as with the alanyl peptides, the ROA changes between H_2O and D_2O indicate that the NH_3^+ or N-H groups may also make a contribution.

1400-1200 cm^{-1}

Unlike Ala_2, no normal mode analysis or isotopomer studies of Val_2 have been done and (as far as I can ascertain) there is no previous Raman data on aqueous Val_2 (although Naik (1992) has recently published Raman spectra of L,D-valanyl peptides in $CHCl_3$ solutions). This makes the assignment of this region difficult because there are more vibrational modes in this region than observed bands.

In Val_2/H_2O a low frequency shoulder appears at 1385 cm^{-1} on the COO^- symmetric stretch band and this vibration could be related to the 1365 cm^{-1} band in Val_2/D_2O. It could possibly be C_α-H^{II} vibrations but the highest Ala_2 C_α-H bend occurs at 1355 cm^{-1} and all other Val_2 C_α-H bends occur at lower frequencies. It may be the shoulder from the 1385 cm^{-1} band in the H_2O sample which relates the COO^- vibration to the N-H/N-D exchange as seen in Ala_2. (Note that the negative ROA feature moves with it.) In Ala_3 Qain and co-workers (1991) find a normal mode at 1400 cm^{-1} which consists of CH_3 symmetric and N-H in plane bends.

There is a band at 1347 cm^{-1} in the deuterated sample which gives very little ROA (it may be slightly positive). In the H_2O sample this band moves to a slightly lower frequency, 1343 cm^{-1}, and acquires a small positive ROA signal. Although the C_α-H deformations occur at 1357 cm^{-1} in valine, this Val_2 vibration does not show the large positive ROA band found with the other C_α-H deformation. Therefore this band is tentatively assigned to the symmetric methyl deformations.

In order to examine the C-H bends properly we consider the D_2O spectra first which, because of the decoupling of the amide III vibration, allows us to examine the C-H bends in isolation. In the deuterated sample the largest ROA signals are associated with the Raman bands at ~1305 and 1325 cm^{-1}. This positive/negative couplet is very similar to, although lower in frequency than, the couplet found in the deuterated Ala_2 spectra (where it occurred at 1330 and 1345 cm^{-1}) and so the bands are assigned to C_α-H bends.

The low frequency 1245 and 1270 cm^{-1} bands correspond to the C-H bends found in valine. In the assignment of valine it is unclear whether these C-H bends orginate in the C_α or C_β, but evidence from other amino acid samples indicates that they are C_β-H bends. These bands give negative ROA (in Ala_2 the low frequency C_α-H vibrations gave positive ROA) and so we make a tentative assignment to C_β-H vibrations, which give a comparatively large negative C.I.D. (Δ) value in valine (Gargaro et al., 1992).

In Val_2/H_2O Raman and ROA bands appear at 1290 and 1305 cm^{-1}. Fasman and co-workers (1978) proposed that these bands were sidechain vibrations. However, our deuteration experiments do not support this, with a large change, particularly of the 1290 cm^{-1} band, in the Raman intensity and the ROA signal upon deuteration. The 1290 cm^{-1} Raman and negative ROA band may be derived from the C_α-H bend which contributed to the

1325 cm^{-1} band in deuterated Val_2. Upon coupling with the N-H in plane deformation it moved to a lower frequency. Naik (1991) assigns a 1288 cm^{-1} signal to an amide III vibration.

The large positive ROA signal which appears at 1303 cm^{-1} in Val_2/H_2O drops to 1300 cm^{-1} upon deuteration. A similar, large positive ROA signal was observed with Ala_2, at the higher frequency of 1345 cm^{-1}; however in Ala_2 this band is far more sensitive to deuteration, dropping by 20 cm^{-1}. It would appear as if the Ala_2 and Val_2 C_α-H deformations not only contribute at different frequencies, but also mix to different extents in the amide III mode.

In the hydrated sample we see that a large amide III band appears at 1245 cm^{-1} with negative ROA, similar to the band found in Ala_2 at 1280 cm^{-1}. It is unclear wheither this band is mixed with the C_β-H bends. Naik's data (1992) shows an amide III band at 1230 cm^{-1}, which could be equivalent to this one.

1200-1050 cm^{-1}

It is in this region of the peptide spectra that we find backbone skeletal stretch modes and methyl and NH_3^+ rocks. In the alanyl series it was difficult to extricate the various components, but hopefully the Val_2 and Ser_3 samples will illuminate the situation.

In the Raman spectrum of the Val_2/H_2O there are two bands that can easily be seen, one at 1115 cm^{-1} and the other at 1165 cm^{-1}. However, the shape of the band envelope indicates that there is another band in between these two that probably occurs around about 1150 cm^{-1}. The 1115 cm^{-1} vibration gives a strong positive ROA signal and the 1150 and 1165 cm^{-1} bands correspond to a strong negative signal.

In the deuterated sample the Raman spectrum shows three bands quite clearly at 1120, 1140 and 1180 cm^{-1}. The deuterated ROA shows that the lower frequency positive side of the Val_2/H_2O couplet vanishes, leaving only a small positive residue. The strong negative ROA band remains at 1150 cm^{-1}. This positive/negative couplet in Val_2/H_2O is very similar to that found in Ala_2/H_2O, where it was assigned to backbone skeletal stretch modes, methyl rocks and small NH_3^+ contributions. In Ala_2/D_2O the positive component of the couplet could be found and the negative component fell by 15 cm^{-1}; neither of these occur with Val_2/D_2O.

In valine a 1110 cm^{-1} band, assigned to antisymmetric CCN stretch, gives a positive ROA band in all the amino acids studied (Gargaro, 1991). The 1140 cm^{-1} (negative ROA at 1150 cm^{-1}) band in valine is assigned to methyl rock and C-C stretch. It is reasonable to carry these assignments to Val_2, particularly in light of the very similar Ala_2 results. However, we can begin to expand on the assignments. In Ala_2 we noted a deuteration effect on the high frequency negative component of the couplet, due to the amino terminal. However we see that as the methyl group contribution increases this band becomes more stable - this confirms our methyl rock assignments. It is difficult to examine the effects on the Raman spectra because D_2O and ND_3^+ deformations both give bands at 1180 cm^{-1}. Raman spectra of solid valine peptides indicate a strong band at 1175-1180 cm^{-1} which remains invariant upon deuteration (Naik, 1992).

When we examine the low frequency side of the couplet we see that the Raman band at ~1115 cm^{-1} remains, but the associated ROA vanishes. Although a similar situation occurred in Ala_2, we could see where the positive ROA features ended up, but in Val_2 the positive ROA intensity does not appear elsewhere. In Ala_2 we assigned this feature to CH_3 rocks with small NH_3^+ contributions. Fasman and co-workers (1978) record that

the terminal NH_3^+ rocks occur at 1140 and 1120 cm^{-1}. Here the CH_3 rocks seem to dominate the Raman spectrum, whereas NH_3^+ rocks dominate the ROA spectrum.

In both the hydrated and deuterated samples we find a band at 1060 cm^{-1}, which gives slightly positive ROA in both cases. In valine a similar band is seen but it is assigned to NH_3^+ rock which cannot be the case here. In serine and cysteine C-N stretch, which gives positive ROA, is also found in this region. However, no similar band is found in valine and so this may indicate a C-N stretch from the peptide link.

1050-700 cm^{-1}

This region of the ROA spectrum contains some very prominent features: certainly there is larger ROA in this region in Val_2 than in Ala_2. The analysis is complicated by the fact that there are no clear ROA signals, but rather several sharp signals which seem to be piled on top of each other. On the experimental side it should be mentioned that in the alanyl samples this region is very prone to strong artifacts.

The region as a whole varies quite considerably upon deuteration, a lot more so than the Ala_2 samples. Although this complicates our analysis, it is not suprising. In their normal mode analysis of Val-$(Gly)_2$ Bandehar and Krimm (1988) note that the ND_3^+ rock combines with a number of other valine side chain C-C stretches in this region.

Unfortunately a study of the Raman and ROA signals from the Val_2 samples in this region lead to, at best, very speculative assignments. This problem is is compounded by the lack of other experimental data on Val_2 molecules and so this region will be discussed no further.

300-700 cm^{-1}

In the low frequency vibrations the problems of assignment become so acute I have not even attempted assignments. Previous experience of low frequency ROA has shown that normal mode analysis of the specific species being studied is essential. In this region we would expect to find the C=O in plane bend and the C=O out of plane bend (the amide IV and VI modes respectively). However, these vibrational coordinates can be the major contributors to bands that are quite widely dispersed and the coordinates will mix very differently depending on main chain conformation and side chain structure. However, this structural dependence, if it can be overcome, can be a great advantage. The following quote is taken from Bandekar and Krimm (1988);

"..it (is) clear that we cannot expect to find characteristic bands for the CO and skeletal deformation modes in the polypeptide chain. These coordinates contribute to normal modes in ways specific to the structures in question. In fact, this means that this region should be particularly sensitive to conformation, and normal mode analysis should provide a useful approach to the determination of structural differences."

4.4 ASSIGNMENT OF THE TRISERINE RAMAN AND ROA SPECTRA.

Tri-serine (Ser_3) is another readily available and soluble peptide. The serine structure is small enough to be amenable to normal mode calculations, and this has greatly helped our assignments. The structure of Ser_3 is shown in Figure 4.1. One interesting feature about the Ser_3 is that it contrasts with the alanyl and valinyl studies because there are no methyl contributions in the vibrational spectra.

The experimental method is described in section 4.2. The Raman and ROA spectra of Ser_3 in 1M HCl and DCl are shown in Figure 4.4. The Raman and ROA of the 1100-1450 cm^{-1} region are shown in Figure 4.5.

1750-1400 cm^{-1}

At 1721 cm^{-1} in HCl and 1729 cm^{-1} in DCl we find the C=O stretch of the COOH and COOD groups, respectively, correlating with the alanyl peptide result. At 1694 cm^{-1} in HCl we find the amide I band with positive ROA, similar to, but with a higher frequency than, Ala_3. The amide I' band is at 1680 cm^{-1} in DCl and the ROA in this region is unclear.

In the HCl we see two shoulders on the low frequency side of the amide I band, one at 1662 cm^{-1} and another at 1569 cm^{-1}. Although the ROA at these frequencies is zero there is a small negative peak inbetween them at ~1600 cm^{-1}. We assign these Raman bands to NH_3^+ antisymmetric bend and amide II, respectively.

The 1471 cm^{-1} band in HCl is assigned to a CH_2 bend, and is analogous to the CH_3 asymmetric bend in Ala_3. It shows no ROA and so is an isolated vibration. Upon deuteration we find that this band mixes with the amide II' vibration and acquires a strong negative/positive couplet. The absence of ROA in the CH_2 bend was noted in serine (Gargaro, 1991; Gargaro et al., 1992), but it was also observed that the different interactions of the CH_2 bend in cysteine produced large positive ROA. We have the same situation here where the amide II' vibration couples with the CH_2 bend to produce a set of interactions that give a large couplet.

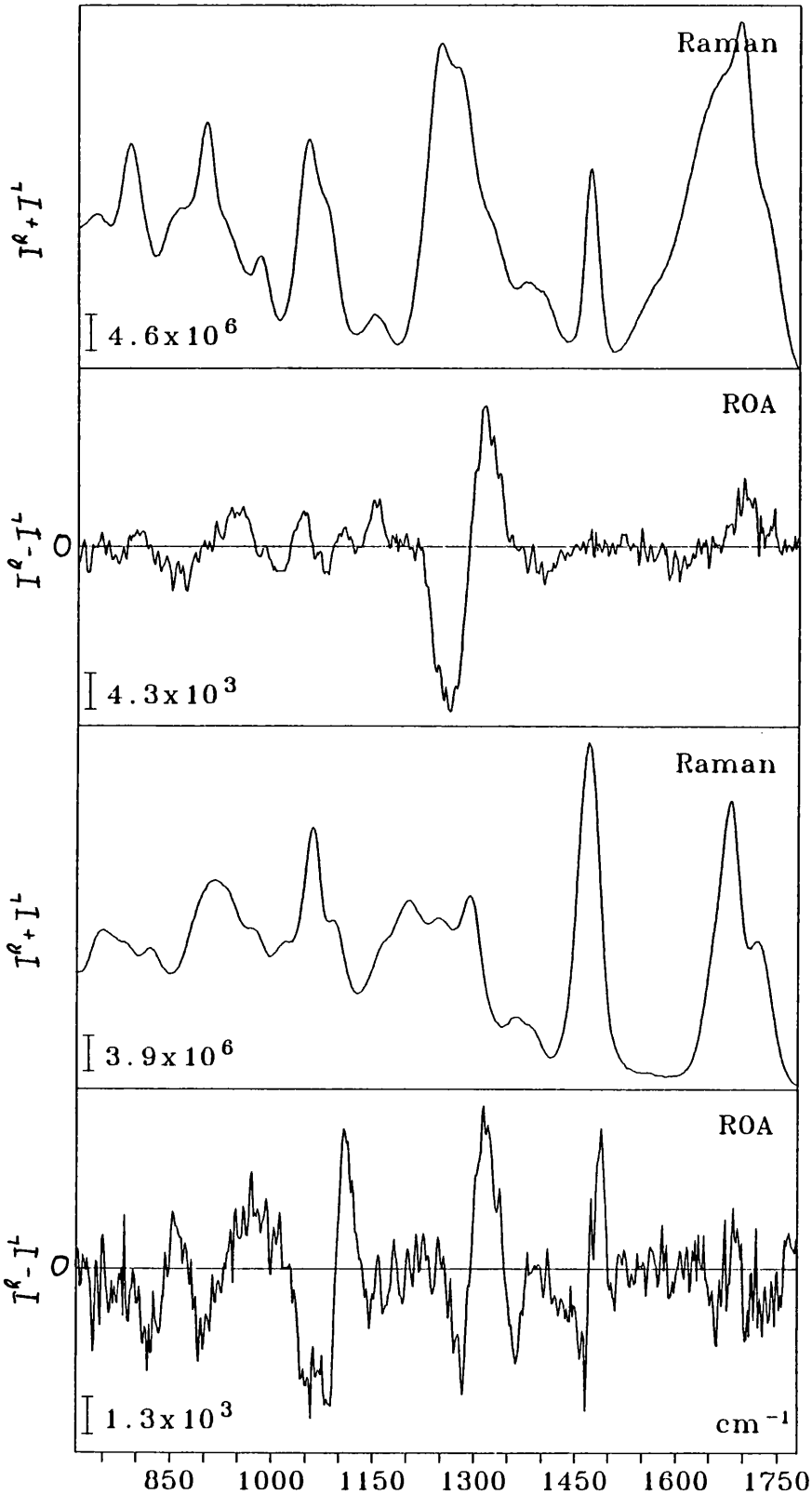


Figure 4.4: The Raman and ROA spectra of Ser_3 in 1M HCl (top) and 1M DCl (bottom).

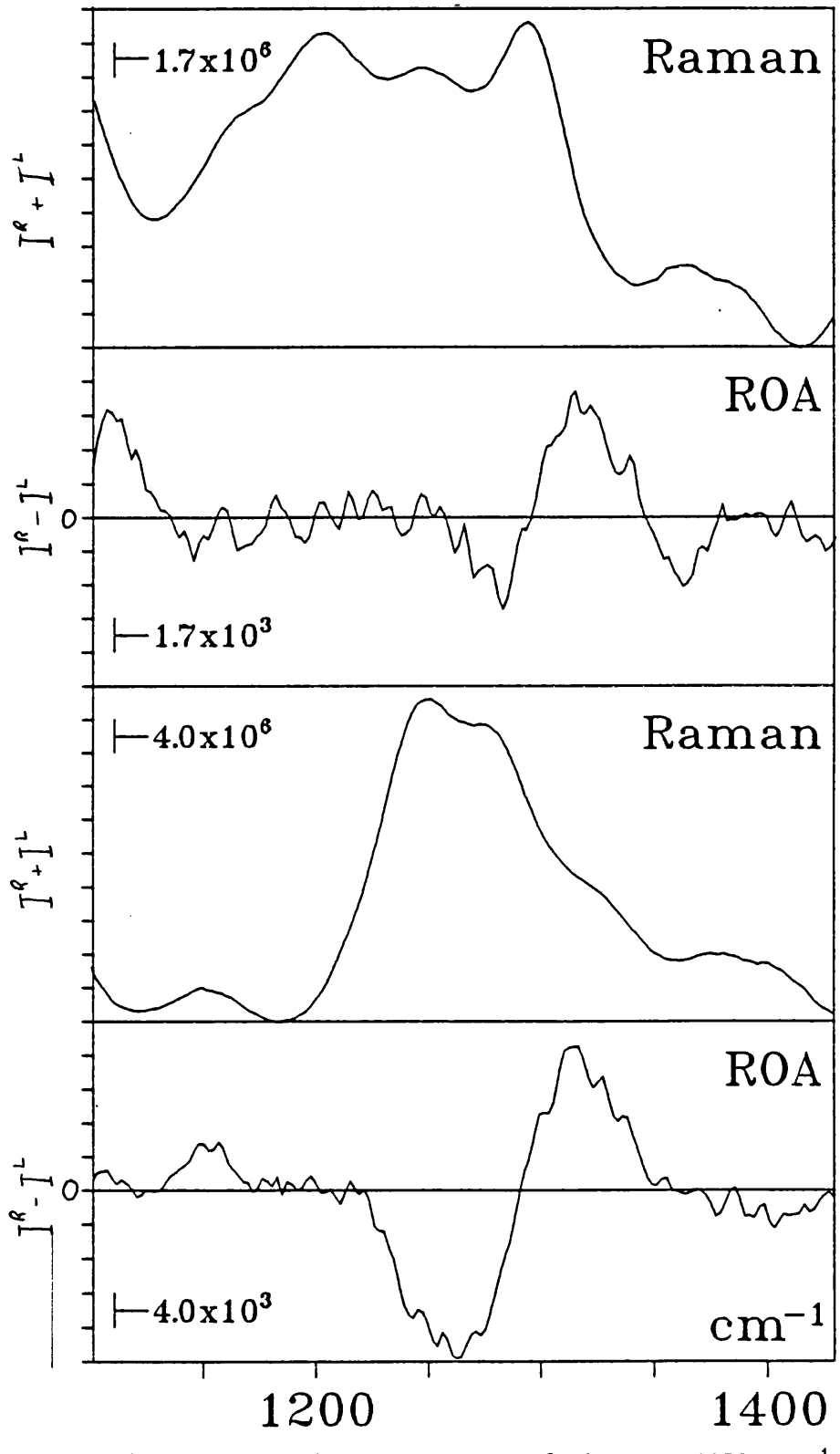


Figure 4.5: The Raman and ROA spectra of the 1100-1450 cm^{-1} region of Ser_3 in 1M DCl (top) and 1M HCl (bottom).

1400-1100 cm^{-1}

In the HCl sample we see two small Raman bands at 1400 and 1375 cm^{-1} and a small negative ROA band at ~1400 cm^{-1}. Upon deuteration these Raman bands fall by 15 cm^{-1} and the ROA band becomes a sharper negative feature at 1360 cm^{-1}. A similar ROA dependence is seen in Ala_3. The deuterium dependence of this region has been observed by Raman (Machida et al., 1979; Gargaro, 1991; Gargaro et al., 1992). It may be that the N-H in plane vibrations control the coupling between the C_α-H bends and the CH_2 bend, and when the coupling mechanism breaks down upon deuteration the ROA signal moves down to the C_α-H bend exclusively. These vibrations at 1400 and 1375 cm^{-1} are likely to be the serine equivalents of the CH_3 bends, probably CH_2 bend. However, Susi and co-workers assign these bands, in solid zwitterionic serine, to highly coupled C-C, COO^-, C-H and CH_2 modes.

In DCl we can see the positive/negative couplet between 1300 and 1380 cm^{-1}. This is the same couplet that is seen in the deuterated samples of Val_2 and the alanyl peptides.

In the 1340-1160 cm^{-1} region we see what appears to be quite significant changes in the Raman and ROA spectrum between the hydrated and deuterated states. However, when this region is considered in isolation (see Figure 4.5) we see that with respect to the 1360-1400 cm^{-1} bands, there is not such a significant change in the Raman and ROA as it originally appeared. More precisely, in HCl the main Raman band occurs at 1250 cm^{-1} with two shoulders at higher frequency, 1274 and 1320 cm^{-1}. The former two bands give strong negative ROA, while the latter gives a strong positive ROA signal. The three Raman bands become four Raman bands, at 1173, 1203, 1251 and 1294 cm^{-1}, in DCl. None of the first three

give any significant ROA, but the last one gives a negative/positive couplet with a very large positive component.

In HCl the strong 1250 cm^{-1} Raman band can be assigned to CH_2 wagging with COH in plane deformations. Although in serine the equivalent Raman band was assigned to COH in plane deformations with a CH_2 wag, in Ser_3 deuteration only reduces the intensity of this band by 20%. This 20% may be accounted for by the collapse of the amide III^1-equivalent that occurs in all the other peptides. However, when we consider the ROA of this band we see that the very strong negative peak vanishes in DCl. This maybe due to one, or both, of two possibilities; either the collapsing of the amide III^1-equivalent that has carried negative ROA in all the other peptides, or the decoupling of the CH_2 and COH vibrations which together give very strong negative ROA. By considering the magnitudes of the signals it appears as if the latter possibility is the most likely. A qualitative glance reveals that in Ser_3 we see a far larger negative signal at this frequency than in Ala_3. The CID (Δ) values show that the intensity of this negative feature is closer to the CH_2/COH vibration in serine than the amide III^A in Ala_3;

Ser_3/HCl	$-5x10^{-4}$
Ser/H_2O	$-6x10^{-4}$
Ala_3/HCl	$-3x10^{-4}$

The Raman bands at 1173 and 1203 cm^{-1} in DCl, which give almost no ROA, can be assigned to the ND_3^+ deformations and a solvent band, respectively.

The two Raman bands in HCl (at 1274 and 1320 cm^{-1}) seem to coalesce into a single band at 1294 cm^{-1} in DCl. Presumably, these represent the observable amide III vibrations. When we look at the ROA we see a negative peak at 1280 cm^{-1}, which possibly formed part of the large

negative signal in HCl. This appears in none of the alanyl DCl spectra and is specific to Ser_3. A good candidate for assignment is CH_2 twist which gives negative ROA in serine and is invariant upon deuteration (Machida et al., 1979).

With a large part of the negative 1220-1350 cm^{-1} ROA couplet assigned to CH_2OH vibrations we see again the ROA we see the 'classic' large positive ROA band associated with the amide III vibrations. If, as we have proposed, the negative 1280 cm^{-1} ROA band in DCl also occurs in HCl then the amide III ROA feature remains exactly the same upon deuteration, even though the Raman bands alter. In Ala_2, Ala_4 and Ala_5 we can see very distinct ROA changes upon deuteration, but in Ala_3 the ROA features in HCl and DCl are the most similar.

This feature causes us to question the amide III coupling situation in Ser_3. Although the nature of the amide III coupling may be such that it generates ROA which minimizes the dependence it has on the N-H in plane bends, we can see from the Raman that the amide III coupling occurs. This illustrates the obvious risk in making deductions from a single molecule, but it also demonstrates the advantage of being able to study a complete series of molecules such as the alanyl peptide series. It may be that in Ser_2 and Ser_4 we would see more distinct changes than we do here. The implications of the similarity between Ser_3 and Ala_3 are discussed further in section 4.5.

In all the previous peptides we have found an exchange-invariant negative ROA band at 1150 cm^{-1}. There is no such band in Ser_3 and this confirms our assignment of it to methyl rock in the other peptides. The 1150 cm^{-1} Raman band in HCl can be assigned to NH_3^+ rock, which gives positive ROA in serine. However, other components are involved as well

because, in the DCl these other components move down by 40 to 1110 cm^{-1} taking the positive ROA with them.

1100-700 cm^{-1}

The assignments in this region become very difficult and specific assignments would only be speculative. Part of the problem lies in extrapolating assignments from the serine monomer, in its zwitterionic form, to the tripeptide, in its cationic form.

The 1052 and 1076 cm^{-1} Raman bands in HCl, which appear to become the 1059 and 1094 cm^{-1} bands upon deuteration, are probably backbone skeletal stretch vibrations. The vibration at 1014 cm^{-1} in DCl belongs to the N-D in plane deformations. Further down between 820 and 980 cm^{-1} in HCl we see the broad negative/positive couplet that we see in the alanyl peptides. Upon deuteration this couplet moves up by 40 cm^{-1}. In this region the coupling of the ND_3^+ vibrations and COD bends will produce significant changes in the Raman and the ROA.

4.5 PEPTIDE CONFORMATION

4.5.1 Introduction

All through the peptide analysis I have avoided the topic of peptide conformation, preferring instead to deal with it in it's own section. Although, as this section will show, ROA is conformational sensitive we are not at the stage of being able to predict aqueous solution conformation from ROA data. This will have to be done by comparing experimental and ab initio ROA experiments or by building up a database of ROA features correlated with known conformational states.

4.5.2 Alanyl peptides

In the alanyl series we see two main changes in the ROA spectra on changing from neutral to alkaline pH: these occur in the amide I region, and in the backbone stretch region between 1200 and 1000 cm^{-1}. There is little doubt that the amide I' (and amide I in Ala_2) move to a lower frequency by 20-25 cm^{-1} upon deprotonation and this could indicate a conformational change.

If we examine the alterations in the backbone stretch region we arrive at conflicting conclusions. The observed changes could be caused by either the NH_3^+ to NH_2 transition changing the normal modes of the vibration alone or a conformational alteration arising from the same deprotonation. A conformational change is implicated by the fact that (although the spectral quality is quite poor in this region of the Ala_4/NaOH sample) the ROA transition can definitely be seen in all the hydrated alanyl peptide spectra. Although the ROA changes are observed in the deuterated Ala_2 spectra, the same transition is not seen in the deuterated Ala_3 or Ala_4 samples, implying that the effect diminishes as the peptide length increases and that it is a change in the vibrational normal modes alone. At the moment there is no explanation as to why this ROA changes.

The amide I' effect is more apparent because it is seen in the more readily accessible Raman spectra. The changes in the amide I' spectra are most likely produced by a spatial interaction between the peptide group and the amino terminal group. The possibility of vibrational contribution from the amino terminal to the amide I' vibration can be discarded because the decrease in frequency is equivalent for both the amide I and I' bands, it occurs in Ala_4 and Ala_5, and normal mode calculations on Ala_3 show that the amide I vibration has no NH_3^+ contributions (Qain et al., 1991). VCD experiments by Zuk and co-workers (1989) lead them to propose that

peptide C=O to amino NH_3^+ hydrogen bonds were present in aqueous peptides. Although we cannot challenge the VCD results, the final conclusion is not satisfactory. The presence of a five membered ring held together by hydrogen bonding in aqueous solution involves an unusual, bent, and probably strained, hydrogen bond. The carboxylic, amine and peptide groups are far more likely to form straighter, and more stable hydrogen bonds with the surrounding solvent water molecules. However, their observation of some form of through space C=O-NH_3^+ interaction bears with our own observation of the amide I band.

The amide III ROA changes between acidic and neutral solutions, but remains very similar between neutral and alkali solutions, which contrasts with the backbone and amide I regions. One possibility is that a conformational change occurs between neutral and alkaline solvents which does not affect the amide III ROA signals: the Φ-dependence model of amide III ROA (see chapter 8) would allow such a possibility. Interestingly enough the peptide conformers proposed by Zuk and co-workers (1989) indicate that between neutral and acidic pH a significant change in the Φ angle of the C-terminal peptide occurs. These conformers are shown in Figure 4.6.

Himmler and Eysel (1989) observed intermolecular hydrophobic clustering of peptides in solution. This implies that attention must be paid to the concentrations used in the various different peptide experiments. Zuk and co-workers used peptide concentrations of ~1.3M, and the ROA spectra of Ala_2 were collected at a concentration of ~3M. Although the intermolecular clustering effects were observed in a similar concentration, 0.3M to 1.7M (Himmler and Eysel, 1989), the VCD and ROA data both appear to support a conformational change between neutral and low pH. No concentration dependent ROA or VCD studies were carried out, and so

the possibility of the observed spectral changes, or the conformational change being due to an intermolecular interaction cannot be ruled out.

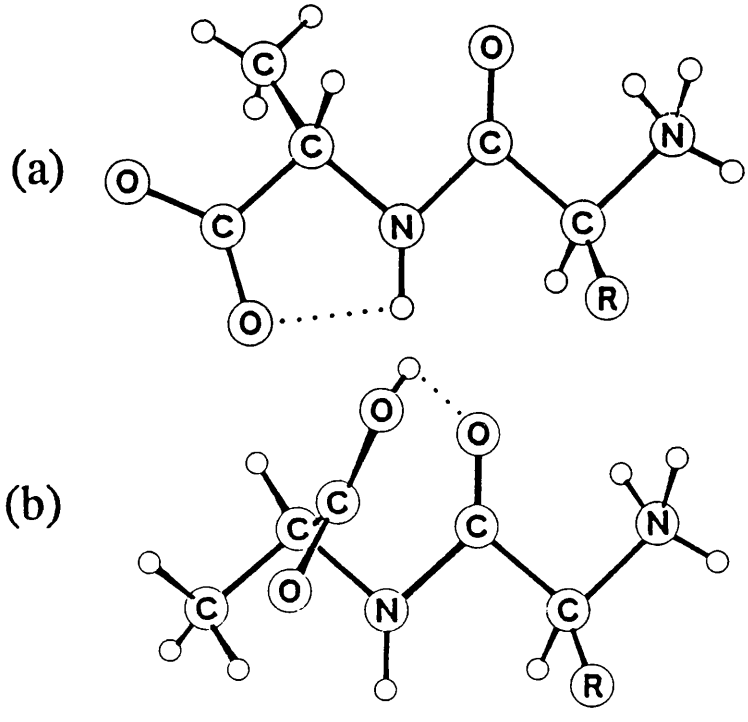


Figure 4.6: Peptide conformers at neutral pH (a) and low pH (b), as proposed by Zuk et al. (1989).

4.5.3 Val_2 and Ser_3

As stated in the overview of the Val_2 results, we see that the Val_2 ROA spectra are sharper and more distinctive than Ala_2 or Ser_3. We propose that this maybe due to hydrophobic clustering of valine residues and such an effect has already been observed by NMR (Wright et al., 1988). It also is interesting to note an effect seen by Himmler and Eysel (1989) in simple amino acids. The authors note that the CO_2^- bending modes show a shift from anisotropic to isotropic scattering as amino acid concentration, or side chain size, increase. This effect they attributed to the formation of micelles (i.e. the clustering of hydrophobic and hydrophilic regions of the molecule).

When we look at the deuterated Val_2 spectra we see a sharp positive/negative amide III couplet between 1285 and 1340 cm^{-1} very similar to the one we see in the alanyl series. However, when we replace the hydrogen and allow the full amide III modes to form, we see a very different ROA envelope: in Ala_2/H_2O the main ROA feature is a large positive band (between 1310 and 1370 cm^{-1}; in Val_2 the identifiable amide III feature is a negative/positive couplet (between 1270 and 1320 cm^{-1}) with a bias to the positive. The different frequency of the couplet may be a function of the conformation but it is more likely that it is a function of the heavier sidechain. This is evident from two features of the peptides that we have studied. Firstly, we know that isotopic substitution will shift ROA features considerable distances (as in the case of the 1100-1150 cm^{-1} region of Ala_2) and we extrapolate this fact to sidechain substitution. Secondly, the deuterated Val_2 ROA spectra are equivalent to the deuterated Ala_2 ROA spectra, so we know that the sidechains will not influence the C_α-H bend ROA envelope, but will alter the frequency. The conclusion is that Val_2 amide III ROA shows large differences compared with Ala_2.

The Ser_3 results show a great resemblence to the Ala_3 results. Again the uncoupled C_α-H bends give a positive/negative couplet. Upon hydrogenation the ROA that appears and can be assigned as amide III ROA is very similar in frequency and characteristics to that found in Ala_3. The hydrated and deuterated Ser_3 spectra look very similar to the hydrated and deuterated denatured hen egg white lysozyme (HEWL), and this may be simple coincidence (with the large negative 1260 cm^{-1} ROA band being CH_2OH vibrations, as suggested above), or a genuine concurrance (where the 1260 cm^{-1} ROA band represents and dynamic amide link, see chapter 8). However, the systematic comparision of various peptides leads to the former conclusion, where as the comparison of two very different molecules, such as Ser_3 and HEWL, is more tenuous.

The amide III ROA features lead to the conclusion that the conformation of Val_2 is not similar to that of Ala_2, but that of Ser_3 is similar to that of Ala_3.

4.5.4 Discussion

The conclusion from the peptide ROA spectra is that the conformation of the alanyl and seryl peptides are similar, but different to that of valanyl peptides. Previous molecular mechanics experiments come to similar conclusions.

Molecular modelling calculations on Pro-Xaa peptides (where Xaa represents a range of amino acids including alanine, valine and serine) concludes that although the inter-residue interations had a major role in determining conformation in the hydrated state, when a solvent sphere was added the water molecules weakened these interactions, levelling out the difference in conformational behaviour (Han and Kang, 1993). However, the occupation of various conformational states can still differ, as is shown in Figure 4.7, which shows that alanine and serine residues occupy a similar range of conformatonal states whereas valine localizes itself predominately in the 'C-state'.

For these peptides the conformation will be decided by a combination of hydrophobicity and steric bulk. The question is how hydrophobic are the alanine residues, and is their hydrophobicity enough to drive them to the conformation induced by the very hydrophobic, and very bulky valine residues? The experimental data which shows the relative solution structures of these three peptides would suggest that the answer is no. However, more importantly we can draw the conclusion from these results that the amide III ROA is conformationally sensitive.

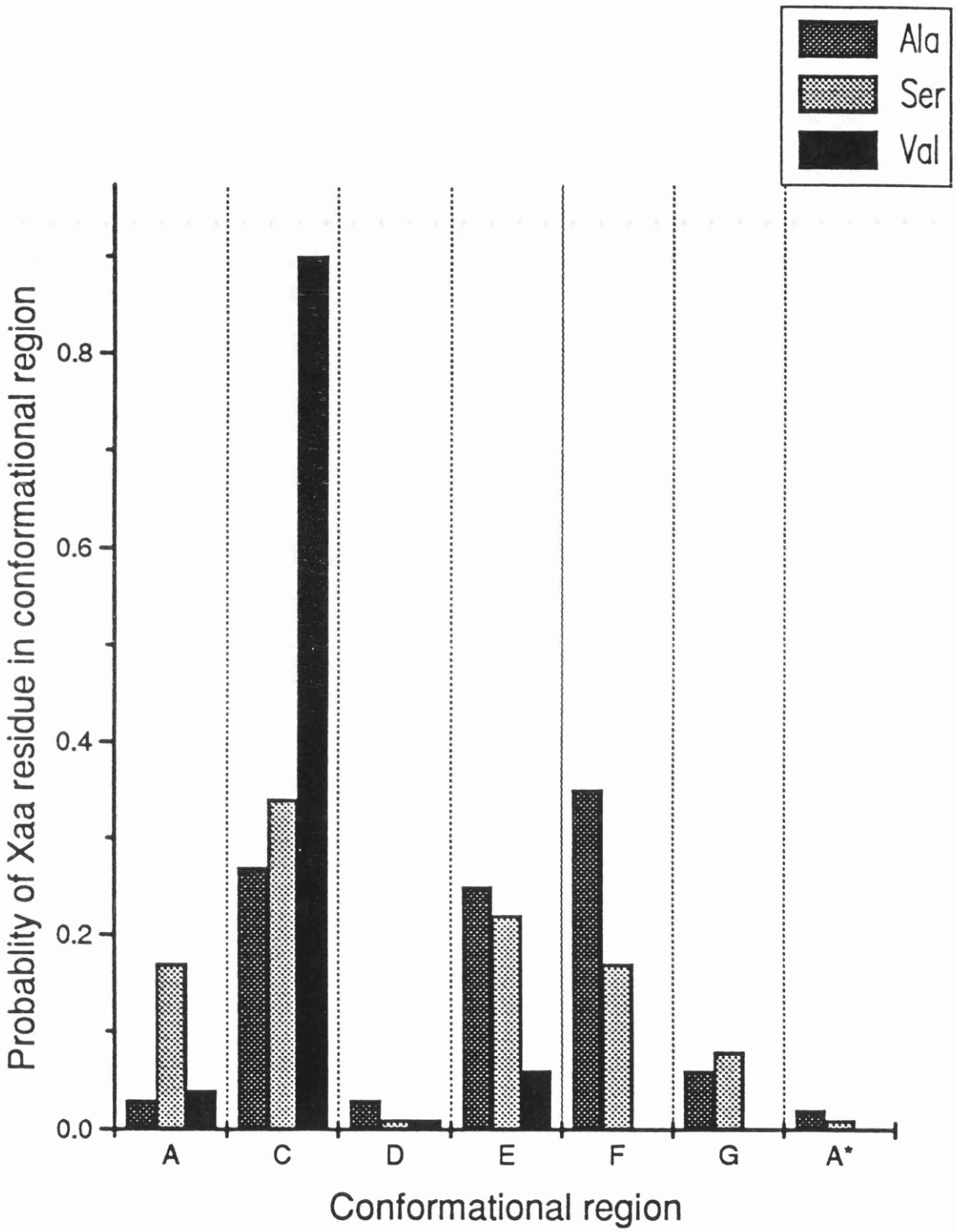


Figure 4.7: The probabilities of occurrence of Xaa residue in Pro-Xaa dipeptides in various conformational regions in the hydrated state. The conformational states are defined in Zimmerman et al., 1977.

4.6 CONCLUSION

In this chapter we have combined the analysis of the Ser_3 and Val_2 ROA spectra with the alanyl peptide assignments in chapter 3. This extra peptide data has:

- reinforced some of the peptide ROA assignments made in chapter 3 (particularly in the 1000-1200 cm^{-1} range).

- shown that the C_α-H deformations appear to give a similar ROA envelope in all the deuterated peptides.

- allowed us to make comparisons between the conformations of Ser_3, Val_2 and the alanyl peptides.

- shown that the amide III ROA is conformationally sensitive, although probably to the Φ angle more than the Ψ angle (see chapter 8).

Chapter 5

Alanine dipeptide

5.1 INTRODUCTION

5.1.1 Background

After examining the poorly understood Val_2 and Ser_3, we turned to a well studied peptide; alanine dipeptide, or N-acetylalanyl-N'-methylamide (see Figure 5.1). Alanine dipeptide is the simplest chiral molecule with two amide links and because of its small size it serves as an appropriate model for theoretical studies of peptide backbone conformation. It has been the subject of several recent ab initio quantum chemical calculations (Rommel-Mohle and Hofmann, 1993; Gould and Hillier, 1993) and molecular mechanics simulations (Smith et al., 1993; Tobias and Brooks, 1992) which describe the conformational behaviour of the peptide under various conditions. Alanine dipeptide has also been studied by several experimental techniques including Raman and IR (Avignon et al., 1973), UVCD and NMR (Madison and Kopple, 1980) and normal mode analysis (Cheam, 1993).

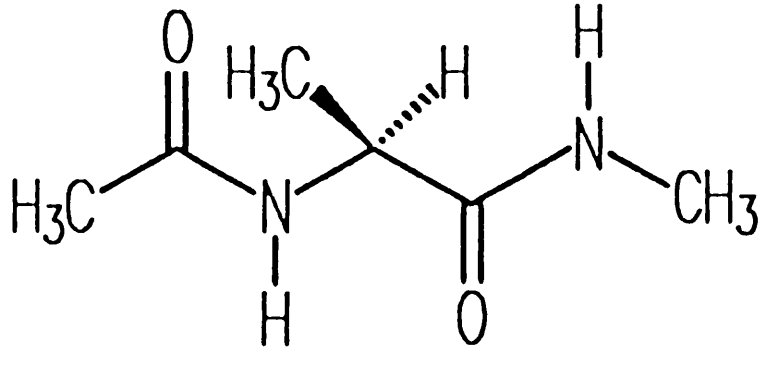


Figure 5.1: The structure of alanine dipeptide.

5.1.2 Peptide Conformation

Alanine dipeptide can adopt several different conformations which are outlined in Table 5.1. It should be noted that different authors have slightly different Ψ/Φ definitions for the same conformer depending upon the stable conformations that their calculations produce.

Table 5.1: The conformations adopted by alanine dipeptide and their appropriate Ψ and Φ angles.

Conformer	$\Psi(^{o})$	$\Phi(^{o})$	Analogous polymer and protein structure.
C_{7eq} [a]	80	-80	Left handed γ turn
C_{7ax} [b]	-80	60	-
C_5 [a]	150	-150	β sheet
α_L [b]	60	60	Left handed α helix
α_R [a,b]	-80	-60	Right handed α helix
P_{II} [a] (β [b])	120	-80	Polyproline helix

a = Madison and Kopple, 1980
b = Tobias and Brooks, 1992

Most of the recent theoretical papers try to establish the effect of water solvation on peptide conformation by comparing the conformation of alanine dipeptide in the gas phase (in vacuo) and aqueous solution phase. Some of these authors have drawn parallels between the conformations found in the gas phase and non-polar solvents (Gould and Hillier, 1993; Rommel-Mohle and Hofmann, 1993). This allows us to use the theoretical calculations of conformation in different 'solvents' to analyse the experimental ROA study of alanine dipeptide in water and chloroform.

The theoretical studies all show that aqueous solvation has a profound effect on the conformational equilibrium of alanine dipeptide. (There is some debate as to whether water induces a different low energy

conformation or not.) Water acts as a 'solvent flattener' and levels out the potential energy surface of the molecule; this lowers the potential energy barriers between the different conformers and increases the conformational freedom of the molecule (Tobias and Brooks, 1992; Rommel-Mohle and Hofmann, 1993). In aqueous solution the rotation of the Ψ and Φ angles are independent (Smith et al., 1993). Minimum energy calculations produce the following results: Tobias and Brooks (1992) predict that the gas phase conformation is P_{II}, with P_{II} remaining and the α_R conformer appearing in the solution phase; Gould and Hillier (1993) make a similar prediction for the aqueous solution conformers, but propose that the gas phase conformers are C_{7eq} and C_S.

The molecular dynamics (MD) simulations of Rommel-Mohle and Hofmann (1993) also show some interesting results. In vacuo the alanine dipeptide fluctuates around two regions: the first region encompasses the C_{7eq}, C_S and P_{II} conformations; the second region is the C_{7ax} conformation which exists as a sharp minimum. Upon solvation the dynamic behaviour of the molecule changes; as before the dipeptide occupies two regions, but this time they are linked. The first region (C_{7eq}, C_S and P_{II}) remains populated but the distribution of of conformational space within it changes (see Figure 5.2); the second region changes drastically with the C_{7ax} conformation dissappearing completely and the α_R conformer becoming predominant (this occurs because the hydrogen bond required to maintain the C_{7ax} conformation is no longer stable). The MD trajectories for in vacuo and aqueous phase results are shown in Figure 5.2.

NMR and UVCD data on the various solution conformations of alanine dipeptide suggests that in non-polar solvents the C_{7eq} conformer

predominates which upon hydration converts to the P_{II} and α_R conformers (Madison and Kopple, 1980).

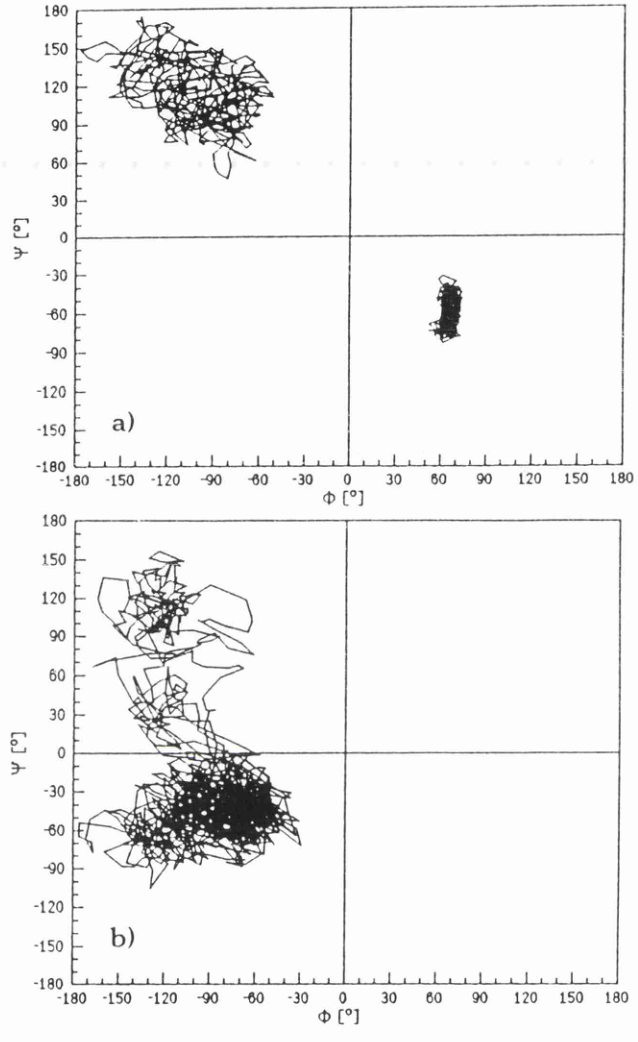


Figure 5.2: The MD trajectories over Ramachandran space for in vacuo and aqueous phase results. (Taken from Rommel-Mohle and Hofmann, 1993)

The experimental and theoretical results appears to agree upon the fact that the α_R conformer is absent in nonpolar solvents, but contributes significantly in aqueous solution. Although there appears to be disagreement on the predominant conformer(s) in nonpolar solvents, most authors agree that the molecule occupies the classical 'β sheet' region of the Ramachandran plot.

5.2 EXPERIMENTAL METHODS

The alanine dipeptide was generously donated by Prof. Prasad Polavarapu of Vanderbilt University, Nashville, Tennessee.

The alanine dipeptide was dissolved in H_2O; the final sample concentration being ~1.8M. The sample was slightly discoloured and activated charcoal was added to the solution to clarify it. The D_2O solvated sample was prepared the same way, except that the alanine dipeptide was lyophilized from D_2O prior to solvation. The sample concentration in D_2O was ~2.1M. The Raman and ROA spectra had to be obtained in two separate sections (as with the Ala_2 and Val_3 spectra).

The alanine dipeptide was purified before being dissolved in $CHCl_3$ by lyophilizing the aqueous fraction of an alanine dipeptide/H_2O/CCl_4 mixture. The lyophilized powder was dissolved in $CHCl_3$ with the final concentration being ~0.7M. This solution was filtered into the ROA sample cell using a Millipore FH (0.5 μm) filter.

Chloroform obviously contains very intense solvent Raman bands which make the acquisition of ROA from the weak alanine dipeptide solution very difficult. These solvent Raman bands not only mask the alanine dipeptide ROA underneath them, but can also distort the surrounding ROA signals by causing strong artifacts. In order to try and minimize these problems and maintain an ideal exposure time the ROA spectrometer was recalibrated so to 'cut off' the strong solvent bands at the edge of the CCD. The Raman and ROA of the alanine dipeptide in chloroform extends from 830 to 1180 cm^{-1} and 1225 to 1780 cm^{-1}. The baseline corrections of the ROA spectra, the Raman solvent subtracted spectra and the determination of the Raman peak frequencies are all done using applications within the LabCalc software package.

5.3 ASSIGNMENTS OF THE ALANINE DIPEPTIDE RAMAN AND ROA SPECTRA.

The Raman and ROA spectra of alanine dipeptide in H_2O, D_2O and $CHCl_3$ are shown in Figure 5.3, 5.4 and 5.5, respectively.

5.3.1 Overview.

The Raman and ROA spectra of alanine dipeptide in H_2O and D_2O nicely illustrate the principle that an ROA signal is only produced by normal mode vibrations that encompass a chiral unit; consequently, the Raman spectrum will have far more bands than the ROA spectrum (23 and 14 respectively for alanine dipeptide in H_2O).

The Raman and ROA spectra of alanine dipeptide in H_2O and $CHCl_3$ are significantly different and this reflects the change in molecular conformation between the two solvents.

5.3.2: Alanine dipeptide in H_2O and D_2O

The Raman spectra of alanine dipeptide in H_2O and D_2O agree with the published spectra (Avignon et al., 1973).

1700-1400 cm^{-1}

The amide I Raman band occurs at 1654 cm^{-1}, but drops by 12 cm^{-1} upon deuteration, showing a significant contribution from the N-H vibration (in Ala_2 the equivalent drop is only 5 cm^{-1}). The amide I and amide I' ROA are both small positive features and appear at the same frequency as their Raman counterparts. The amide II vibration occurs at 1562 cm^{-1} and gives a positive ROA feature. (Unlike the alanyl peptides, NH_3^+ or COO^- vibrations cannot occur in this region.) This bears out the assignments of similar Raman and ROA bands in the alanyl peptides to the

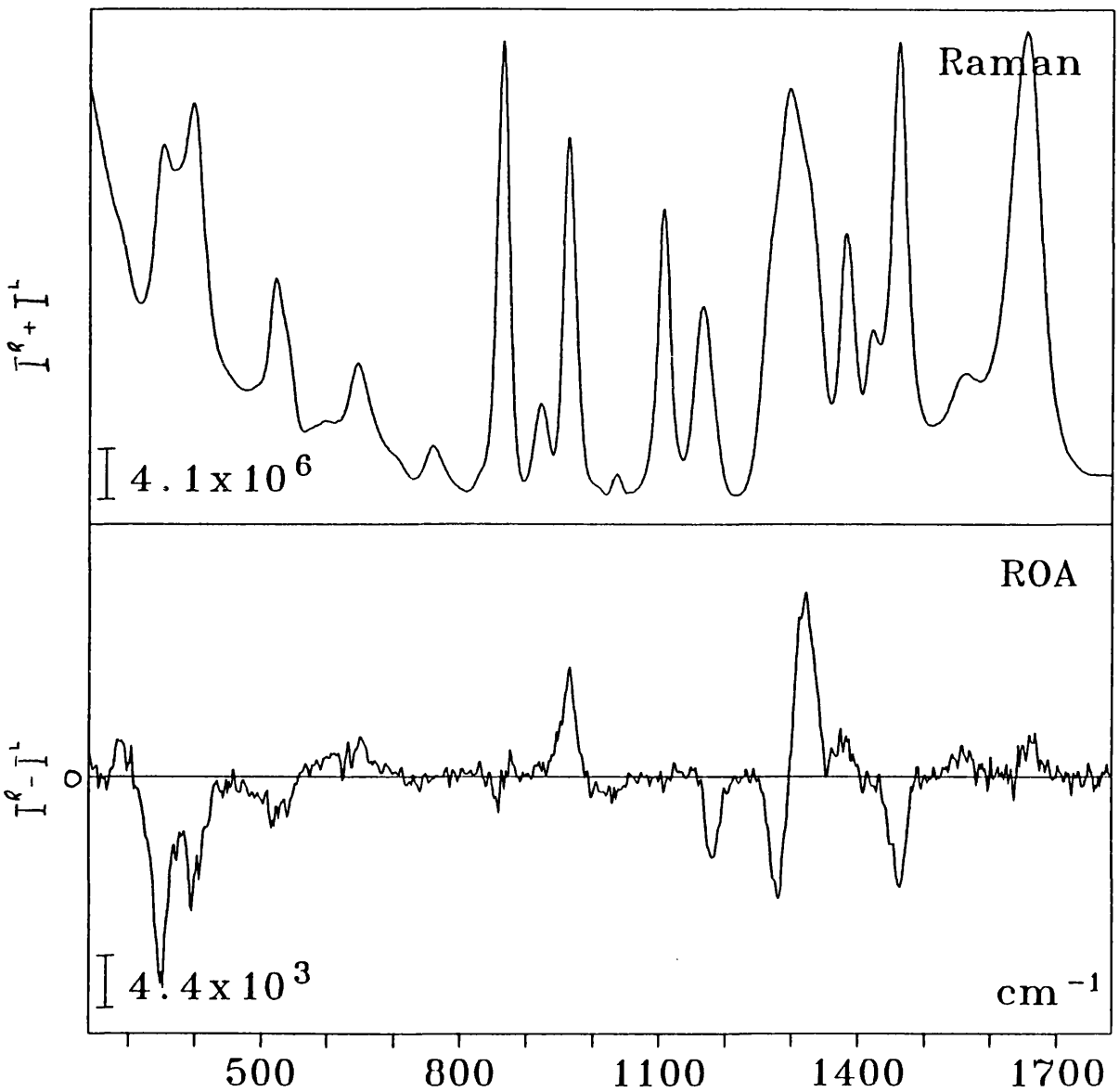


Figure 5.3: The Raman and ROA spectra of alanine dipeptide in H_2O

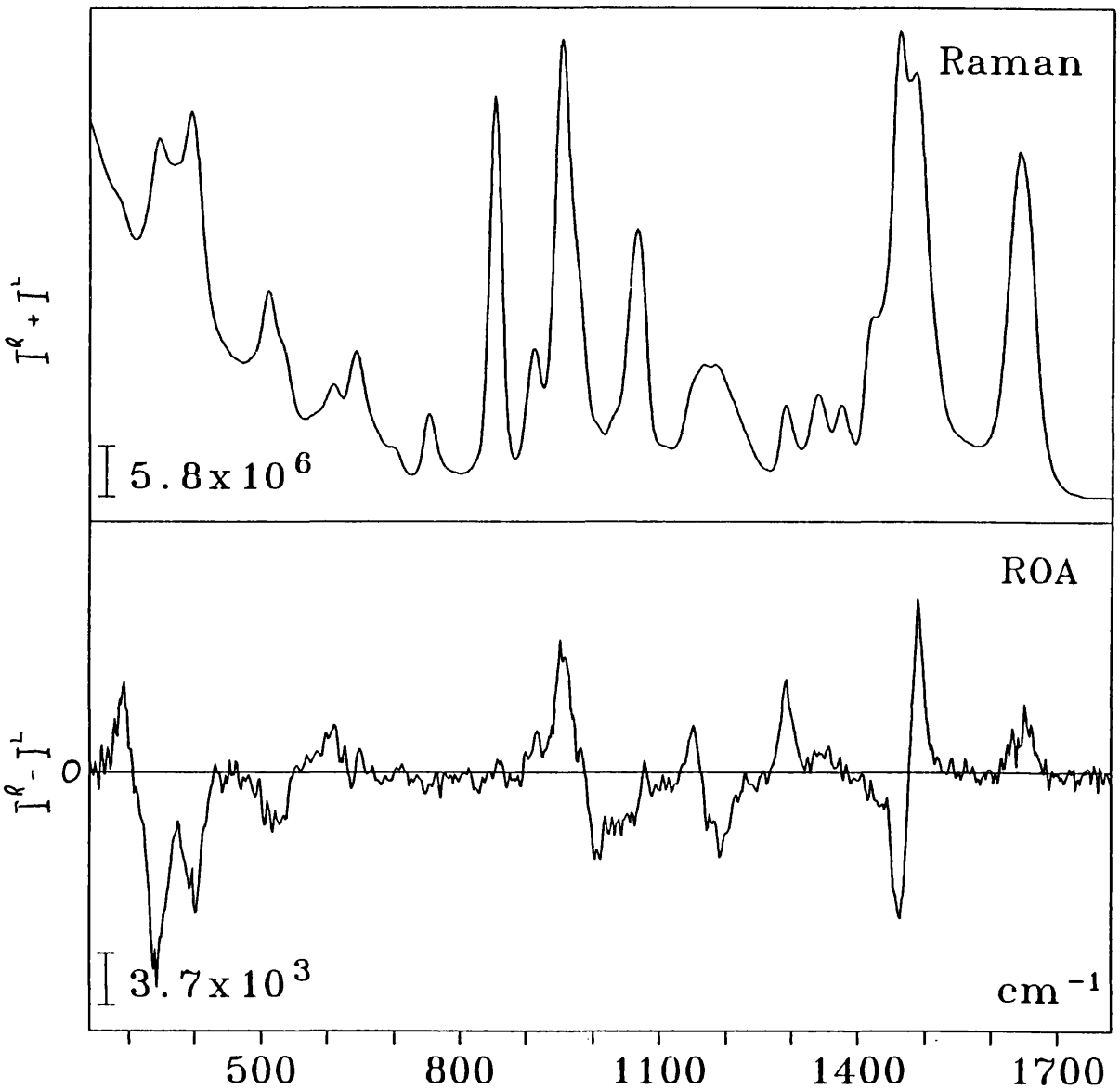


Figure 5.4: The Raman and ROA spectra of alanine dipeptide in D_2O.

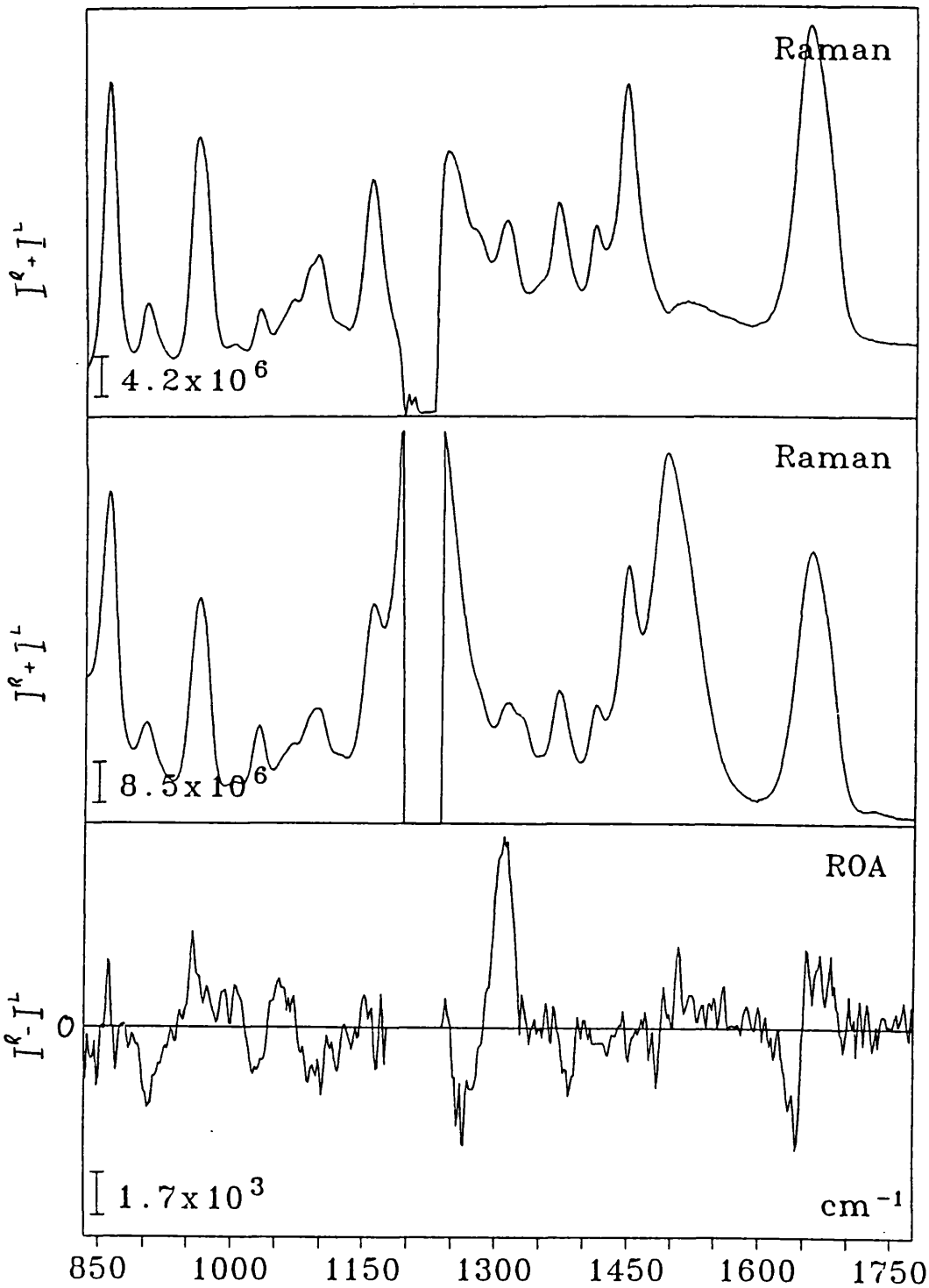


Figure 5.5: The solvent subtracted Raman spectrum (top), the Raman (middle), and the ROA (bottom) of alanine dipeptide in $CHCl_3$.

amide II vibration. The amide II' vibration occurs at 1484 cm^{-1}, where it mixes with the asymmetric methyl deformations as borne out by the new negative/positive ROA couplet. The signs of this couplet match the signs of the original uncoupled vibrations (from the hydrated sample) in the same way as they did for Ala_2.

All three asymmetric methyl deformations appear in the strong Raman band at 1461 cm^{-1} (Cheam, 1993). These vibrations give a strong negative ROA signal contrasting with alanine alone where the asymmetric methyl deformation gives positive ROA (Barron et al., 1991). One possible reason for this may be the additional terminal-blocking methyl groups coupling with alanine methyl group, similar to the asymmetric methyl deformation coupling found in ab initio ROA calculations on twisted, chiral N-methylacetamide (Polavarapu et al.,1994).

The Raman band at 1423 cm^{-1} (in both hydrated and deuterated samples) belongs to a methyl symmetric deformation but there appears to be confusion over whether this band originates from the C-terminal-blocking (Cheam, 1993) or the N-terminal-blocking methyl group (Avignon et al., 1973). This vibration gives no ROA signal.

1400-1200 cm^{-1}

The deuterated sample shows three Raman bands in this region at 1376, 1340 and 1291 cm^{-1}. The 1376 cm^{-1} vibration moves to a slightly higher frequency, 1383 cm^{-1}, upon hydration and is assigned to symmetric methyl deformations from the alanine methyl group with a small contribution from the C_α-H bend (which may couple with the N-H in plane deformations to give the deuterium sensitivity). This vibration gives a small positive ROA feature in the hydrated alanine dipeptide.

The other two Raman bands in the deuterated sample are assigned to C_α-H deformations; both vibrations give positive ROA signals as they did in alanine (Barron et al., 1991). However with the alanine dipeptide the lower frequency vibration gives a far larger ROA intensity than the higher frequency vibration and this may be due to the extra contributions from amide C-N stretches which the normal mode analysis places in this region (Cheam, 1993).

In the Raman spectrum of the hydrated alanine dipeptide a band appears at 1298 cm^{-1} with high and low frequency shoulders. This represents the amide III vibration(s) and will have several contributions, including C_α-H bends, C-N stretch, N-H in plane bends, but the exact nature of the coupling within these vibrations cannot be deduced without normal mode calculations based on the solution structure of alanine dipeptide. The Raman spectra of solid and solvated alanine dipeptide are significantly different in this region (Avignon et al., 1973). The ROA spectrum shows a large negative/positive couplet which represents the amide III vibration and demonstrates how the presence of the N-H in-plane deformation can alter the ROA of this region.

1200-1000 cm^{-1}

In D_2O two Raman bands appear at 1184 and 1168 cm^{-1}; they are not particularly clear because the D_2O bend vibration raises the Raman baseline at ~1200 cm^{-1}. The low frequency vibration, which gives positive ROA, is assigned to C-N stretch (Avignon et al., 1973), while the high frequency vibration, which gives negative ROA, is assigned to methyl rocks. The ROA signals confirm these assignments; the 1110 cm^{-1} C-N stretch of alanine gives a positive ROA feature (Barron et al., 1991); the methyl rocks from the peptides give negative ROA feature (Ford et al., 1994 and this work).

The Raman band at 1108 cm^{-1} has previously been assigned to the C_α-C stretch and methyl rocks (Cheam, 1993), or C_α-N stretch (Avignon et al., 1973). However, this Raman band has no ROA associated with it, and so a more probable assignment is relatively pure (perhaps terminal) methyl rocks, which the ab initio calculations on N-methylacetamide have shown occur here (Polavarapu et al., 1994). The 1037 cm^{-1} Raman band gives slightly negative ROA and is invariant upon deuteration. This vibration has been assigned to methyl rocks (Cheam, 1993; Avignon et al., 1973) and our observation do not contradict this. Upon deuteration the 1108 cm^{-1} Raman band moves down to a lower frequency of 1067 cm^{-1} and acquires a negative ROA signal. This is probably due to the mixing of the CH_3 rocks with N-D in plane deformations.

1000-800 cm^{-1}

The 963, 921 and 864 cm^{-1} Raman bands in hydrated alanine dipeptide all drop by 9 cm^{-1} upon deuteration. All these bands are assigned to C-C stretches and methyl rocks (Avignon et al., 1973; Cheam, 1993). Only the high frequency vibration carries it's strong positive ROA band (which also drops in frequency with the Raman band upon deuteration). It would appear as if only the high frequency C-C stretch involves the chiral C_α.

800-300 cm^{-1}

The assignments of the low frequency Raman and ROA bands requires accurate normal mode calculations on the solvated alanine dipeptide molecule. However, this region is important for ROA studies. In their ab initio ROA calculations on N-methyl acetamide, Polavarapu et al. (1994) pointed to two ROA signals that were especially sensitive to conformation: methyl bending modes and coupled C=O bending vibrations, which appear

in this region of the spectrum. Possible assignments for this region are given in Table 5.2.

Table 5.2: Possible assignments for the low frequency vibrations of alanine dipeptide.

Raman (cm^{-1})		ROA (cm^{-1})		Assignment
H_2O	D_2O	H_2O	D_2O	
761	755	-	-	C=O opb[a]
700	700	-	-	C=O ipb[b]
645	645	+	?	Amide IV[a,b,c]
596	d	+?	+	C=O opb[a] (Amide VI[c])
521	511	-	-	C=O ipb[a] (Amide IV[c])
396	396			
352	347	-	-	

5.3.3 Alanine dipeptide in chloroform.

1700-1400 cm^{-1}

The amide I frequency of alanine dipeptide in chloroform, 1655 cm^{-1}, is very similar to that of aqueous alanine dipeptide. The amide II vibration is not immediately apparent in the Raman spectrum, but does appears as a broad band in the solvent subtracted spectrum, peaking at 1520 cm^{-1}. Although the band has probably been distorted by solvent subtraction, there is still a substantial drop in amide II frequency when going from water to chloroform. Krimm and Bandekar (1986) have noted that the amide II vibration is conformationally sensitive, with α-helical polypeptides consistently generating high frequency amide II vibrations (1545 - 1555 cm^{-1}) and β-sheet structures generating a broader range of amide II frequencies. This indicates a decrease in the population of the α-helical conformer upon solvation in the $CHCl_3$, which agrees with the other experimental evidence.

The amide I ROA is a negative/positive couplet which differs from the amide I ROA in water. Although this may indicate that the amide I ROA is conformationally sensitive, there are other factors that also have to be considered: Krimm and co-workers have recently (1994) proposed that the amide I vibration is solvent dependent due to coupling between the amide I frequency and the bending mode of H_2O molecules hydrogen bonded to the amide link. In addition to this, it is also known that the nature of the alanine dipeptide hydrogen bonding changes from intermolecular in H_2O to intramolecular in the gas phase. Any ROA from the amide II vibration is not discernible.

The asymmetric and symmetric methyl deformations appear at 1455 and 1415 cm^{-1} in the Raman spectrum. They have no ROA intensity (although any ROA intensity on the low frequency band would have been wiped out by a cosmic ray strike during the spectral acquistion). The frequency of these vibrations drops by up to 8 cm^{-1} in the nonpolar solvent. In addition to this the negative ROA intensity of the asymmetric methyl deformations vanishes in chloroform. This appear to indicate that the coupling between the CH_3 deformations found by Polavarapu and co-workers (1994) confers conformational sensitivity on the asymmetric deformations.

The 1370 cm^{-1} Raman band is assigned to the symmetric methyl deformation from the alanine methyl. It's slight negative ROA contrasts with it's positive ROA signal in aqueous solvent. This concurs with the behaviour of the asymmetric methyl deformation ROA.

The amide III Raman appears to be made up of three separate bands at 1310, 1275 and 1245 cm^{-1}; this differs substantially from the amide III Raman of the alanine dipeptide in water. The ROA, on the other hand, does not change with the frequency and intensity of the negative/positive couplet being very similar between water and chloroform. This

demonstrates that quite different molecular mechanisms are responsible for ROA and Raman. Other ROA signals have shown that a conformational change has occurred and the amide III Raman backs that conclusion up. However, the amide III ROA remains the same in both solvents, and so the conformational change that does occur does not appear to alter that signal.

Does this observation give us any insights into the mechanism that controls the amide III ROA? Even for the Raman amide III the conformational dependence is not clear; the traditional Ψ dependence model (Lord, 1977) has now been called into question with a Φ dependence model now being proposed (Krimm and Bandekar, 1986). A simplistic view of amide III coupling would suggest that any amide III vibration would be more sensitive to Φ angle changes - where the coupling deformations are adjacent to each other - than to Ψ angle changes - where the coupling deformations are separated by a carbonyl group. (This 'Φ-dependence model' of ROA amide III dependence can be used to explain ab initio ROA results and experimental ROA data on the alanyl peptides.) However, in their experimental analysis of alanine dipeptide conformations Madison and Kopple (1980) stated that:

"..the important (conformational) change in going from chloroform to water is probably in the distribution of Φ.".

This leads to the conclusion that the Φ-dependence model does not apply to the experimental ROA spectra of alanine dipeptide. One possible explanation is that the dynamic behaviour of the solvated dipeptide makes the analysis in terms of simple conformational parameters very difficult. Another is that the Φ-dependence model does not apply to the alanine dipeptide molecule because the dipeptide does not contain peptide links between chiral C_α's.

1200-850 cm^{-1}

The methyl rocks and C-N stretch vibration appears at 1165 cm^{-1} in the Raman spectrum and there appears to be no ROA associated with it. This absence of ROA may reflect a change in conformation or may be the result of artifacts caused by the adjacent solvent band. However, ab intio ROA results on alanine dipeptide have shown that the ROA of this vibration is conformationally sensitive (Polavarpu and Deng, 1994).

The next three Raman bands appear at 1100, 1090 and 1070 cm^{-1}. This differs substantially from the single Raman band found in in the aqueous alanine dipeptide. The ROA becomes a positive/negative couplet. The Raman and ROA changes also point to a conformational change.

The 1030 cm^{-1} Raman band, which gives negative ROA, shows the same behaviour as it did in aqueous solvent, where it was assigned to isolated methyl rocks. The following three Raman bands at 965, 900 and 865 cm^{-1} are assigned to C-C stretches and methyl rocks. However, these vibrations do show differences between the two solvents: the high frequency vibration loses some of it's positive ROA intensity; the middle vibration not only drops in frequency by about 20 cm^{-1}, but also acquires negative ROA intensity.

5.4: ALANINE DIPEPTIDE CONFORMATION.

The Raman and ROA spectra of alanine dipeptide in aqueous solvent and chloroform are substantially different, reflecting the change in molecular conformation between the two solvents. However, to determine any further conformational data from these results requires ab initio ROA calculations. Ab initio ROA calculations have been done on 10 alanine dipeptide conformers (Polavarapu, private communication), but only the

following have been published: P_{II} (called C_{7eq}-C_5 by Polavarapo and Deng), α_R, C_5 and α'(where $(\Psi,\Phi)=(-160^o,-40^o)$ (Polavarapu and Deng, 1994). These four ab initio ROA spectra are shown in Figure 5.6 along with the experimental ROA spectra for the H_2O and $CHCl_3$ solutions. Unfortunately no one ab inito ROA spectrum reproduces the experimental ROA features. The experimental ROA above 1350 cm^{-1} shows none of the ab inito ROA features and this has been seen with previous comparisons of other molecules (Barron et al., 1992c). The correct amide III ROA couplet is only reproduced by the α_R and P_{II} conformers. The ROA comparison between 950 and 1350 cm^{-1} appears to indicate that the H_2O conformer is P_{II} and the $CHCl_3$ conformer is α_R (as shown by the connecting lines in Figure 5.6). In contrast, the 850 to 950 cm^{-1} region shows that H_2O conformer could be α', α_R or C_5, and that the $CHCl_3$ conformer is probably P_{II}.

These results are disappointing and no definite conclusions on the solvated conformations of alanine dipeptide can be reached. The most likely reason for the poor comparison between the experimental and theoretical results is the dynamic behaviour of the solvated alanine dipeptide: the current ab initio-experimental comparison relies on being able to deduce the dominant single conformer in solution, rather than determining the conformational distribution of the molecules. A larger range of ab initio calculations will be required to overcome this problem (Polavarapu, private communication). It is interesting to note that Madison and Kopple (1980) could not determine the conformational ratios for the alanine dipeptide in polar solvents. An additional source of confusion may be the poor quality of the ROA spectrum in chloroform.

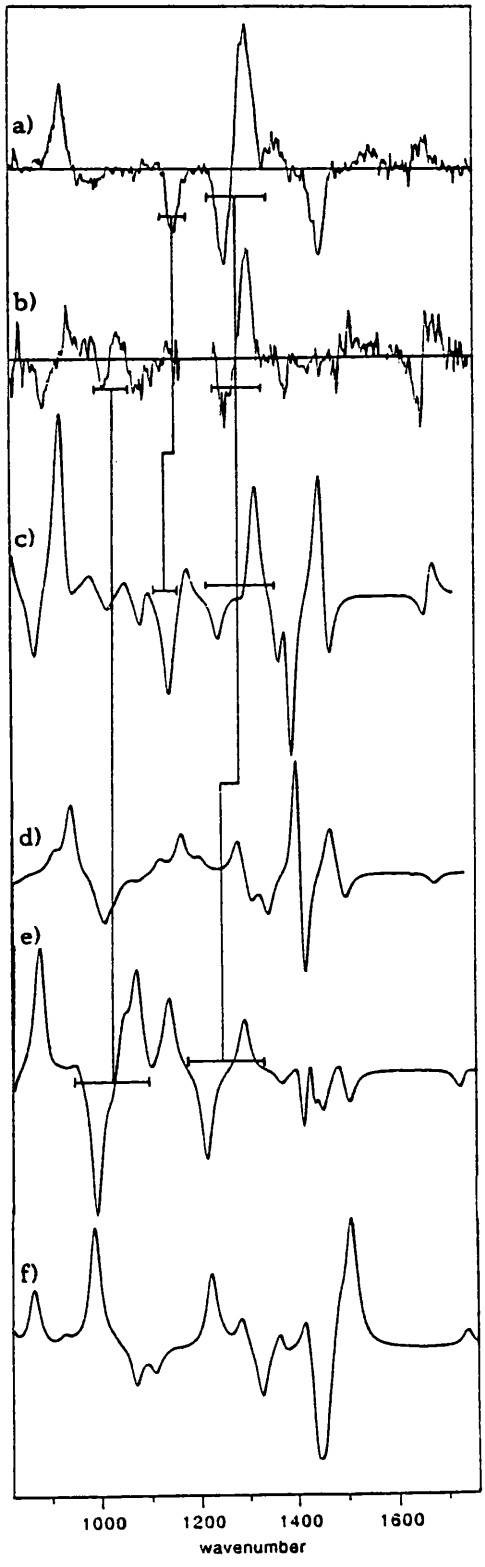


Figure 5.6: Experimental ROA spectra for alanine dipeptide in H_2O (a); in $CHCl_3$ (b); ab initio ROA spectra for alanine dipeptide in the conformers P_{II} (c); α' (d); α_R (e) and C_5 (f). (Adapted from Polavarapu and Deng, 1994)

5.5 CONCLUSION

The Raman and ROA analysis of alanine dipeptide, the smallest and most intensively studied chiral dipeptide has led to:

- comprehensive assignment of the Raman and ROA spectra.
- the first Raman and ROA spectra from an identical peptide in two different conformational distributions.
- interesting comparisons of experimental and ab initio ROA data.

Chapter 6

Hen Egg White Lysozyme

6.1 INTRODUCTION

6.1.1 Background

One of the first proteins we studied in detail was Hen egg white lysozyme (HEWL). In many ways HEWL is an ideal starting protein for a new analytical technique, such as ROA. It's advantages are that it is small, cheap, readily available in high purity and well studied.

HEWL is an anti-bacterial enzyme with a β(1-4) glucoaminidase activity which allows it to hydrolyse the walls of bacterial cells. There are at least three different kinds of lysozyme and HEWL belongs to the c-type lysozyme family which encompasses all the mammalian lysozymes, moth lysozymes and some bird lysozymes. Plant, fungus and bacterial lysozymes are also known. HEWL is 129 amino acids long and has a molecular weight of 14315 amu. Its sequence is shown in Figure 6.1.

```
                                          10
Lys-Val-Phe-Gly-Arg-Cys-Glu-Leu-Ala-Ala-Ala-Met-Lys-Arg-His-

                  20                                      30
Gly-Leu-Asp-Asn-Tyr-Arg-Gly-Try-Ser-Leu-Gly-Asn-Trp-Val-Cys-

                                          40
Ala-Ala-Lys-Phe-Glu-Ser-Asn-Phe-Asn-Thr-Gln-Ala-Thr-Asn-Arg-

                  50                                      60
Asn-Thr-Asp-Gly-Ser-Thr-Asp-Tyr-Gly-Ile-Leu-Gln-Ile-Asn-Ser-

                                          70
Arg-Trp-Trp-Cys-Asn-Asp-Gly-Arg-Thr-Pro-Gly-Ser-Arg-Asn-Leu-

                  80                                      90
Cys-Asn-Ile-Pro-Cys-Ser-Ala-Leu-Leu-Ser-Ser-Asp-Ile-Thr-Ala-

                                          100
Ser-Val-Asn-Cys-Ala-Lys-Lys-Ile-Val-Ser-Asp-Gly-Asn-Gly-Met-

                  110                                     120
Asn-Ala-Trp-Val-Ala-Trp-Arg-Asn-Arg-Cys-Lys-Gly-Thr-Asp-Val-

Gln-Ala-Trp-Ile-Arg-Gly-Cys-Arg-Leu
```

Figure 6.1: The amino acid (primary) sequence of HEWL.

There are two proposed mechanisms by which HEWL hydrolyses saccharide oligomers. In both mechanisms glycosidic cleavage occurs between the D and E sugars (see section 6.1.3 for definitions). The original mechanism was developed by Blake, Johnson, Phillips and co-workers between 1965 and 1972, and is based on X-ray data and model building studies. This mechanism, Scheme I in Figure 6.2, proposes that Glu-35 donates a hydrogen to the glycosidic oxygen with the resulting cation being stabilized by Asp-52; the addition of water finishes the process. The second scheme was suggested by Post and Karplus in 1986 to explain additional evidence that had been observed in the preceeding 10 years. Scheme II does not require the distortion of sugar D and involves the cleavage of the endocyclic C^1-O^5 bond. (For a more detailed description of HEWL, and these two mechanisms, see the recent review by McKenzie and White, 1991)

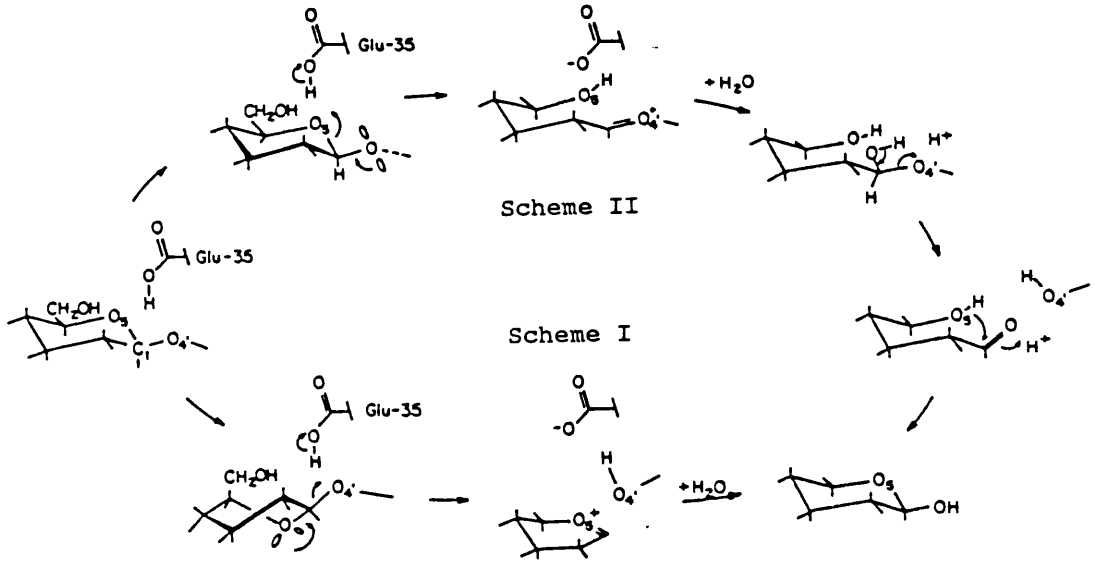


Figure 6.2: The two proposed mechanisms for HEWL hydrolysis. (Adapted from McKenzie and White, 1991.)

The structure and dynamics of enzyme-substrate complexes is one of the most important areas of biophysical chemistry. Fortunately, the enzyme-substate complex of HEWL is well understood, and so one of the most logical experiments to do is HEWL inhibitor studies. The classical HEWL inhibitors and substrates are oligomers of N-acetyl-glucosamine (NAG_n). Short chain saccharides (n=1 to 3) act as inhibitors, long chain saccharides (n=4 to 6) act as enzyme substrates. NAG, NAG_2 and NAG_3 are readily available from Sigma, and the interactions between HEWL and NAG oligomers are well studied. Figure 6.3 shows the structural formula of NAG_3.

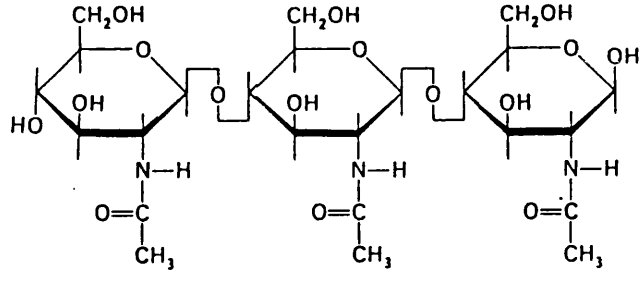


Figure 6.3: The structural formula of NAG_3.

6.1.2 Crystal structures of HEWL

The crystal structure of HEWL, in both its free and inhibitor-bound form, was elucidated in 1965 (it was the first X-ray crystal structure of an enzyme) and since then the structure has been examined in various space groups and at higher resolution. The structure of HEWL is shown in Figure 6.4.

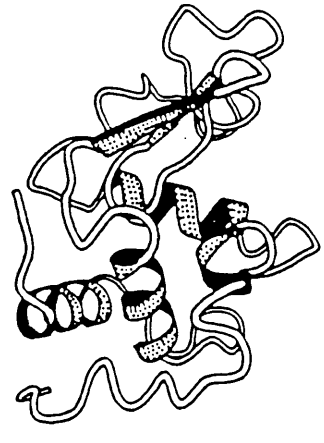


Figure 6.4: The structure of HEWL. (Adapted from McKenzie and White, 1991)

HEWL is approximately 40 x 30 x 30 Å across, with a deep cleft on one side. (The substrate binds in the cleft, which gives it access to the buried catalytic residues.) The cleft defines two lobes; the α-helix lobe, which consist of residues 1-39 and 85-129, and the β-sheet lobe, which consists of residues 40-84. All eight cysteines in HEWL form disulphide bridges; 6-127, 30- 115, 64-80 and 76-94. The secondary structure elements are given in Table 6.1.

Table 6.1: The secondary structural elements in HEWL.

Structure	Residues
α-helix A	4-15
B	24-36
C	88-99
D	108-115
3_{10} helix	79-84
β-sheet	2-3, 39-40 (2 stranded)
	42-46,50-54,57-60 (3 stranded)

The residues that line the cleft include Glu 35, Thr 43, Asn 46, Asp 52, Leu 56, Glu 57, Ile 58, Asn 59, Trp 62, Trp 63, Arg 73, Ile 98, Asp 101, Gly 102, Asn 103, Ala 107, Trp 108, Val 109 and Ala 110. Between the active site cleft, with its flexible residues, and the rigid hydrophobic core lies a collection of residues known as the hydrophobic box. The hydrophobic box includes residues such as Trp 28, 108 and 111, Tyr 23 and Met 105. Recent studies have shown that the hydrophobic box plays a role in inhibitor binding and thermal denaturation (Miura et al., 1991; Bernard et al., 1990).

HEWL can be crystallized in four different crystal space groups (triclinic, tetragonal, monoclinic and orthorhombic), but the structure is essentially the same in all four, with the exception of part of the β sheet region and some of the longer sidechains (McKenzie and White, 1991).

6.1.3 Structural changes upon inhibitor binding.

The substrate binding cleft has six sugar binding sites, A to F, and when NAG_3 binds to HEWL it binds to sites A, B and C. Each of the sugars is buried to a different extent with the order of solvent exposure being A>B>C, (Cheetham et al., 1992). The N-acetyl sidechains of the sugar residues in sites A and B are both solvent accessible, and are therefore hydrogen bonded to solvent molecules. The polar side of the sugar bound at the A side (which contains the N-acetyl group) lies towards a mainly hydrophilic region of the cleft, defined by Arg 73 and Asp 101. Cheetham et al. (1992) observed that the apolar face of the sugar in site B was directly above Trp 62, and proposed that this indicated a hydrophobic interaction. Similarly the hydrophobic face of the sugar in site C lies in a hydrophobic environment generated by Ile 98, Ala 107, Trp 108 and Val 109. The hydrogen bonding that occurs between NAG_3 and HEWL is summarized in Table 6.2.

Table 6.2: The hydrogen bonding between NAG_3 and HEWL (Cheetham et al., 1992)

Site	Sugar atom	Protein residue
A	O6	Asn 103
B	O6	Asp 101
C	N2	Ala 107 (amide C=O)
	O7	Asn 59 (amide N-H)
	O3	Trp 63
	O6	Trp 62
	O1	2 x H_2O

When the NAG_3 binds, Trp 62 and Trp 63 move by 1.0 and 0.7 Å, respectively, towards sites B and C, closing around the NAG_3 and maximizing the potential for intermolecular interactions. Similarly Ala 107,

Asp 101, Asn 103 and Trp 108 all undergo slight conformational changes for the same reason. In the crystal structures regions of the protein remote from the active site show small perturbations. A comparison of the temperature factors show that residues involved in binding the NAG_3 stiffen up, while at sites remote from the active site little variation in temperature factors are found (Cheetham et al., 1992).

The β(1-4) glycosidic linkages are twisted in the bound trisaccharide. The linkages have an angle of 28° and 70° between the C and B, and B and A site sugars respectfully. This allows a hydrogen bond to form between O5 of sugar B and the O3' hydroxyl group of sugar C. This hydrogen bond is not found between sugars A and B or in the chitobiose, where the degree of twist about the glycosidic link is 54° (Cheetham et al, 1992).

6.1.4 Solution structure of HEWL.

The NMR assignment of HEWL was virtually completed in 1988 (121 of the 129 residues) by Redfield and Dobson (1988), allowing them to produce a model of solution secondary structure based on Nuclear Overhauser Effect Spectroscopy (NOESY). They concluded that the secondary structure that they found was identical to that found in the crystal structure. The remaining 8 residues were assigned in 1993 (Smith et al., 1993a). This NMR work has been moved on by Dobson's Oxford group, and the NMR experiments have been of great help in drawing conclusions about our own ROA experiments.

NOE measurements of HEWL in solution lead to the conclusion that the main chain fold and the internal sidechain residues have well defined structure that corresponds with the analogous crystal but this was not the case for the surface residues (Smith et al., 1993a). The α-helical and

β-sheet regions of solution HEWL were equivalent to those found in the crystal stucture and the loop region between 61 and 78, and the C-terminal 3_{10} helix (120-124) had well defined structure in the solution model. (The β-sheet and 61-78 loop only show defined stucture when the comparison is made locally; these residues deviate quite largely from the average NMR and crystal structures when the whole molecule is aligned. It is unclear whether this represents a true effect or is just a result of few NOE distance restraints in that part of the molecule). Residues 79-84, which adopt a 3_{10} helix in the crystal structure, are conformationally labile. (However, it is suggested that this is due to insufficient NOE data rather than large scale conformational disorder.) The NMR group could find no evidence for the hinge bending fluctuation postulated by theoretical calculations. Interior side chain residues, which are sheltered from the solvent, have a well defined conformation. Surface residues occupy multiple conformations, as indicated by a lack of NOE data and coupling constants. The active site residues on the other hand are better defined than the other surface residues and are closer to the conformations found in crystal structures. The exception is Trp 62, which occupies multiple conformations.

Early conventional Raman studies which compared the HEWL crystal and solution spectra concluded that the two structures were very similar (Yu and Jo, 1973).

6.1.5 Structural studies on HEWL and NAG_n: NMR data

The Dobson group also produced quite a detailed study on the NMR of HEWL-NAG_n complexes (Lumb et al., 1994). This study compared the chemical shifts of HEWL resonances in both the free and the bound form, and interpreted the results using computational chemical shift calculations derived from the HEWL and HEWL-NAG_n crystal structures. The chemical

shift of a resonance is a summation of various contributions from a variety of sources: accordingly this data is very sensitive to protein conformation but is also difficult to interpret.

The chemical shifts induced in HEWL by NAG, NAG_2 and NAG_3 binding are very similar with only a slight increase in the effect as the oligosaccharide size increases. The largest chemical shifts are found at binding site C, with small shifts being observed in site B for NAG_2 and NAG_3.

Residues further away from the active site are also changed. They include the main chain resonances of Asp 66, Leu 75, Cys 76, Ile 78 and residues in the C-terminal loop from 122-128. The side chain residues of Trp 62, Trp 63 and Trp 108 also change. The chemical shift of the hydrophobic box residue Trp 28 alters because of the change in conformation of Trp 108 (this change was also observed by Fukamizo et al. 1992). However, the predicted chemical shifts of remaining resonances in the hydrophobic box region are not found in this NMR study.

An earlier NMR study (Bernard et al., 1990) used C^{13} NMR and showed completely different results to the Oxford group. They suggested that conformational rearrangement occurred in both the active site and hydrophobic box regions of the protein and that the effects increased proportionally as oligomer size increases from NAG to NAG_3. The three active site residues of Trp 62, Trp 63 and Trp 108 were found to be particularly sensitive to NAG_n binding. They used these changes to try and confirm the conformational and dynamical alterations found for Trp 62 and Trp 63 in crystallographic and earlier NMR studies.

Another recent NMR study by (Kumagai et al., 1993), which concentrated on the binding of NAG oligomers to Trp 62 mutants, has shown that upon

mutation the major changes in the chemical shifts of HEWL occur in the β-sheet region (41-60) and the loop region (61-78). It was proposed that these two regions also play a role in substrate binding.

Overall, there seems to be no consensus of opinion from the NMR results on the actual effects of NAG oligomer binding. The relative effects of NAG binding appear to be in doubt, as do the conformational alterations in the hydrophobic box. However, the changes in the active site, the Trp 62, Trp 63 and Trp 108 residues and the loop residues 70 to 75 seem to be agreed upon.

6.1.6 Amide hydrogen exchange in HEWL.

The NMR data on HEWL has allowed a variety of hydrogen exchange experiments to be done and these have been used to determine HEWL folding mechanisms (Miranker et al., 1991), pH dependence and electrostatic effects in HEWL (Delepierre et al., 1987), the structure of denatured and partially folded states (Radford et al., 1992; Buck et al., 1993) and the factors influencing hydrogen exchange (Pedersen et al., 1991). Consequently HEWL becomes a well studied system for hydrogen exchange and we can apply this information in our ROA studies.

The amide hydrogens in solvated HEWL can be divided into 4 categories, depending on their rate of hydrogen exchange (see section 6.3.5 for more precise details). It has been shown that hydrogen exchange is slow for those amides that are involved in secondary structure or buried beneath the protein surface, but neither factor alone determines the exchange rate. The ROA features of simple peptides are dependent upon the deuterium exchange of the amide link and this implies that for proteins the ROA features will also change when the amide links become deuterated. Consequently, because the NMR data has determined the

hydrogen exchange rates for all the structural features of HEWL we could start to assign ROA bands to HEWL structural features if we reproduced the NMR experiment.

6.1.7 The denatured states of HEWL.

A comparison was made of the exchange rates of amide hydrogens in the denatured state and the native state (Radford et al., 1992). It showed that the pattern of protection (i.e. areas of slow hydrogen exchange) in the denatured three disulphide form of HEWL (CM^{6-127} HEWL) had no relation to the pattern of protection in the native state, and the hydrogen exchange observed were relatively similar to those predicted for a small peptide. However, the majority of the amide links that did show some protection were from hydrophobic residues and it was suggested that this was an indication of solvent exclusion, i.e. hydrophobic clustering. The overall conclusion was that the denatured states of HEWL contain little persistent residual order.

Further data is available on the nature of the denatured states, and was obtained using magnetization transfer experiments. It is possible to assign some of the NMR peaks in denatured states of HEWL by carrying out a magnetization transfer experiment where the native and denatured states are interconverting. This allows the detection of magnetization transfer between different states, so assignments for the native HEWL can be used as a springboard to assign resonances in the denatured HEWL. The NMR spectra of thermally denatured HEWL shows that the residues do not give the chemical shifts that would be expected of a random coil protein. Upfield shifts remain, which imply that hydrophobic clustering occurs between the hydrophobic residues and the aromatic residues in the chain. This clustering occurs between residues that are far apart in the protein sequence. When the disulphide bonds are broken, or a chemical

denaturant is added, the chemical shifts move closer to the values that would be expected for a random coil but there are still notable differences. Most of the interactions that occur in these latter two states are due to interactions between adjacent residues (or residues that are close together in the chain), i.e. Trp 62 and Trp 63 interactions. This is because when the disulphide bonds are broken the protein is able to sample far larger conformational space and interactions between distant parts of the chain will become less likely.

6.2 EXPERIMENTAL METHODS

6.2.1 Sample preparation.

All the ROA protein samples were prepared in the same general manner. The sample preparation and some purification procedures are discussed below.

The HEWL samples all originated from Sigma, and were either Grade I (3x lyophilized, dialysed and crystallized) or Grade VI (3x crystallized). In the earlier experiments (which are not definitive but are included in certain sections to illustrate the reproducibility of the ROA spectrum) Grade VI HEWL was used. In the later experiments Grade I HEWL was used because it appeared to give better Raman and ROA spectra. All the HEWL samples were run in buffer. All protein concentrations were measured by UV spectroscopy, with the extinction coefficient, E_{280nm}, equal to 2.65 $mlmg^{-1}cm^{-1}$ (Cooper, 1974). HEWL enzymic assays were done using the standard Sigma procedure of cell wall hydrolysis of Micrococcus lysodeikticus.

The preparation of all the HEWL samples followed the same standard procedure, which is as follows:

1) All the buffers were made up using Analar grade chemicals and distilled, or purified and deionized, water. For the deuterated buffers D_2O was purchased from Fluorochem. The pH was measured using a Corning 220 pH meter which was calibrated that day. The pH quoted for deuterated solutions are uncorrected pH values, but the pH can be converted to a pD values by adding 0.41 (Covington et al., 1968).

2) The protein solution was made up in a small glass sample tube (capacity ~2ml) which was pointed at the bottom. This tube was cleaned using chromic acid and filtered spectroscopic grade methanol before being rinsed and dried. The protein and buffer were both weighed directly into this tube. Any mixing of the solution that was required was done by hand using a pasteur pipette. Stirring bars and vortexing were avoided.

3) In the later experiments decolorizing charcoal, purchased from Aldrich, was used to purify the sample. It was found that charcoal lowered the fluorescence background of the protein sample, this improved Raman and ROA spectra quality and cut down of sample defluorescence times (see (6)). The amount of charcoal added was between 0.5 and 1 mg for a 0.2 to 0.3 ml protein sample. The charcoal was found to have no effect on the concentration of HEWL or denatured HEWL, but it was found to affect the pH of acetate buffer slightly (it raised the pH of 100mM acetate buffer by 0.04 pH units). The charcoal was dispersed in the protein solution, and removed by centrifugation after 5 or 10 minutes.

4) Before being filtered the protein solution was centrifuged for 10 minutes in a plastic eppendorf tube. The eppendorf tubes were soaked in buffer overnight before use.

5) The sample was transferred, by pasteur pipette, to a glass Hamilton syringe (250 or 500μl). The syringe was washed out with buffer before use and the sample was filtered through a Millipore GV4 (0.22 μm) filter into a quartz microfluorescence cell (this filter was pre-washed with ~1 ml of buffer). The filter lowers the concentration of the protein because it is wet when then protein sample if pushed through it, but this can be partly avoided by discarding the first few drops of filtered protein solution.

6) The microfluorescence cell was centrifuged and placed in the laser beam. It was found that the fluorescence background of a sample would decrease (in an exponential fashion) with time. During this 'burning down' process the ROA quality is improved, but it was also important to allow the sample to stabilize in the beam. This took between 6 and 24 hours depending on the initial quality of the sample.

In the attempt to improve the sample quality two other techniques were tried, Fast Protein Liquid Chromatography (FPLC) and dialysis. The purification by FPLC was not satisfactory, particularly as ROA requires significant amounts of protein. A trial with α-lactalbumin was carried out, but there seem to be no overall improvement in the quality of the Raman spectrum. Dialysis was carried out on some early protein samples, and especially useful was microdialysis (Pierce Microdialyser), which allowed a complete change of buffer salts in 2 hrs. The advantage of dialysis was that it ensured that any salts in the protein sample, which may alter pH, would be removed. However, dialysis increased the complexity of sample preparation, and never produced consistent improvements in Raman quality. For the latter experiments the dialysis procedure was abandoned.

6.2.2 HEWL-NAG_3 experiments.

An HEWL sample was made up and an ROA spectum was obtained. This sample was then taken out of the sample cell and mixed with a two molar excess of NAG_3. The sample was then syringed (without a filter) back into another sample cell and another ROA spectrum was taken. In the third and final sample, extra precautions were taken because literature sources noted that NAG_3 is slowly broken down by HEWL. The HEWL-NAG_3 sample was run within 3-4 days of mixing. There was a slight drop in HEWL concentration between the free and inhibitor bound experiments. This was because the GV4 filter was used to filter the sample. To ensure that HEWL is fully bound a 50, 2 and 1.3 molar excess of NAG, NAG_2 and NAG_3, respectively, are required (Lumb et al., 1994).

The rate of NAG_3 breakdown can be roughly calculated from data in a review by Imoto et al. (1972). They quote the activation energy of NAG_3 breakdown as 16 kcal/mol and the first order rate constant at pH 5 and 40°C as 8.3×10^{-6} s^{-1}. These results mean that under our experimental conditions the half life of NAG_3 is 6 days.

I would like to acknowledge the assistance of Mr. G. Wilson who helped obtain the ROA spectra.

6.2.3 Denatured HEWL

There are three main ways of denaturing HEWL (chemical denaturation, thermal denaturation and disulphide reduction), but not all are suitable for ROA studies. Chemical denaturants such as urea and guanidinium.HCl have the disadvantage that they require very high concentrations (6M) to denature HEWL. The extra Raman bands generated by these chemicals will produce large artifacts in the ROA spectrum (similar to the effect of the

chloroform bands with the alanyl dipeptide - see chapter 5). Sodium dodecyl sulphate (SDS) will denature proteins with a far more modest (0.2M) concentration. However, attempts to use SDS gave protein samples that had a very high fluorescent background.

The difficulties with thermal denaturation lie in the length of time it takes to produce an ROA spectrum. Although it was initially thought that thermal denaturation at the high concentrations used for ROA studies would lead to gelation, this does not occur immediately. However, gelation or hydrolysis would probably occur if the temperature was maintained at 80-90°C for 24 hours. This, combined with the practical difficulties of doing temperature-dependent ROA studies, made us look to the third denaturation technique.

Disulphide reduction in HEWL lowers the stability of the folded state by increasing the entropic contribution of the unfolded state (Cooper et al., 1992). If all four disulphide bonds in HEWL are reduced then the protein unfolds, and measurement of its ROA is possible without the experimental problems present with the other two strategies. Disulphide reduction is reversible and so the free sulphydryl groups are usually carboxymethylated to prevent reoxidation. The experimental procedure for breaking the disulphide bonds was adapted from Creighton, 1989.

1) 6M Gd.HCl was made using 10 mls of 0.1M Tris base.

2) This was titrated with concentrated HCl or glacial acetic acid to give a pH of 8.

3) 500 mg of HEWL was dissolved in the solution and stirred.

4) A 5 molar excess of dithiolthreitol (DTT) was added.

5) The free sulphydryl groups were blocked by a 5 molar excess of sodium iodoacetate, or iodoacetamide, which was added 20 minutes after the DTT.

6) After 20 minutes the pH was lowered to 1 or 2 using acetic acid or HCl.

7) The solution was dialysed against six or seven 3 litre volumes of distilled water, and then lyophilized.

However, the carboxymethylated HEWL was insoluble (only producing a 2% solution in 0.01M NaOH). The amidomethylated HEWL produced a 6% solution in H_2O and 0.1M HCl and precipitated rapidly as concentrated NaCl was added. Neither of these samples were sufficiently concentrated to give a suitable ROA spectrum.

The NMR experiments on denatured HEWL had used disulphide reduction, except they had left the sulphydryl groups unblocked and maintained a low pH to prevent reoxidation (Evans et al., 1991). The above procedure was altered by missing out step (5) and dialysing the reduced HEWL against 0.01M HCl. Although this sample was restricted to low pH solvents, it would dissolve in 10 mM citrate buffer to give an 8% solution. The number of free sulphydryl groups was checked using the Ellman assay (Creighton, 1989). In all cases the HEWL contained between 7 and 8 free sulphydryls per protein molecule. The denatured HEWL ROA spectrum was produced using a 10mM citrate buffer at pH=3.

6.2.4 Deuterated HEWL.

Completely deuterated HEWL samples can only be prepared by denaturing the protein in the presence of D_2O or a deuterated buffer. The completely deuterated HEWL samples were prepared in two separate ways. In the earlier experiments 0.3 mls of a 150 mg/ml HEWL solution was made up using deuterated acetate buffer. This was prepared as for an ROA

experiment, and sealed in a microfluorescence cell. This cell was heated up in a water bath to 80°C for 15 minutes and then gradually cooled. This procedure was tested by doing enzyme assays on HEWL in hydrated acetate buffer. This denaturation procedure appears to have a minimal effect on HEWL activity which only decreases by 8% upon heating. In the later experiments thermal denaturation was used again, but this time the HEWL (100 mg) was dissolved in D_2O (15 g), and left in a drying oven in a stoppered 100 ml round bottomed flask at 60°C for 5 days. After lyophilization it was prepared as a normal HEWL sample. Activity assays on this fully deuterated HEWL showed that it's enzymatic activity was completely retained.

The partially deuterated HEWL samples were made up by dissolving HEWL (50 mg) in deuterated acetate buffer, pH=4.0 (0.25 g). They were then left at room temperature for either 7 days or 25 days, before being refrigerated prior to preparation for ROA analysis.

6.3 ASSIGNMENTS OF THE HEWL RAMAN AND ROA SPECTRA.

6.3.1 Native HEWL

The Raman and ROA spectra of HEWL are shown in Figure 6.5

1750-1500 cm^{-1}

The Raman band at 1665 cm^{-1} is a combination of the amide I vibrations and the solvent band (H_2O bend). On the low frequency side of the amide I band is a small shoulder, at 1625 cm^{-1}, which remains invariant upon deuteration and denaturation. This band originates in aromatic ring stretch vibrations from the phenylalanine, tyrosine and tryptophan residues (Harada and Takeuchi, 1986). The next two bands at 1580 cm^{-1} and 1555 cm^{-1} are

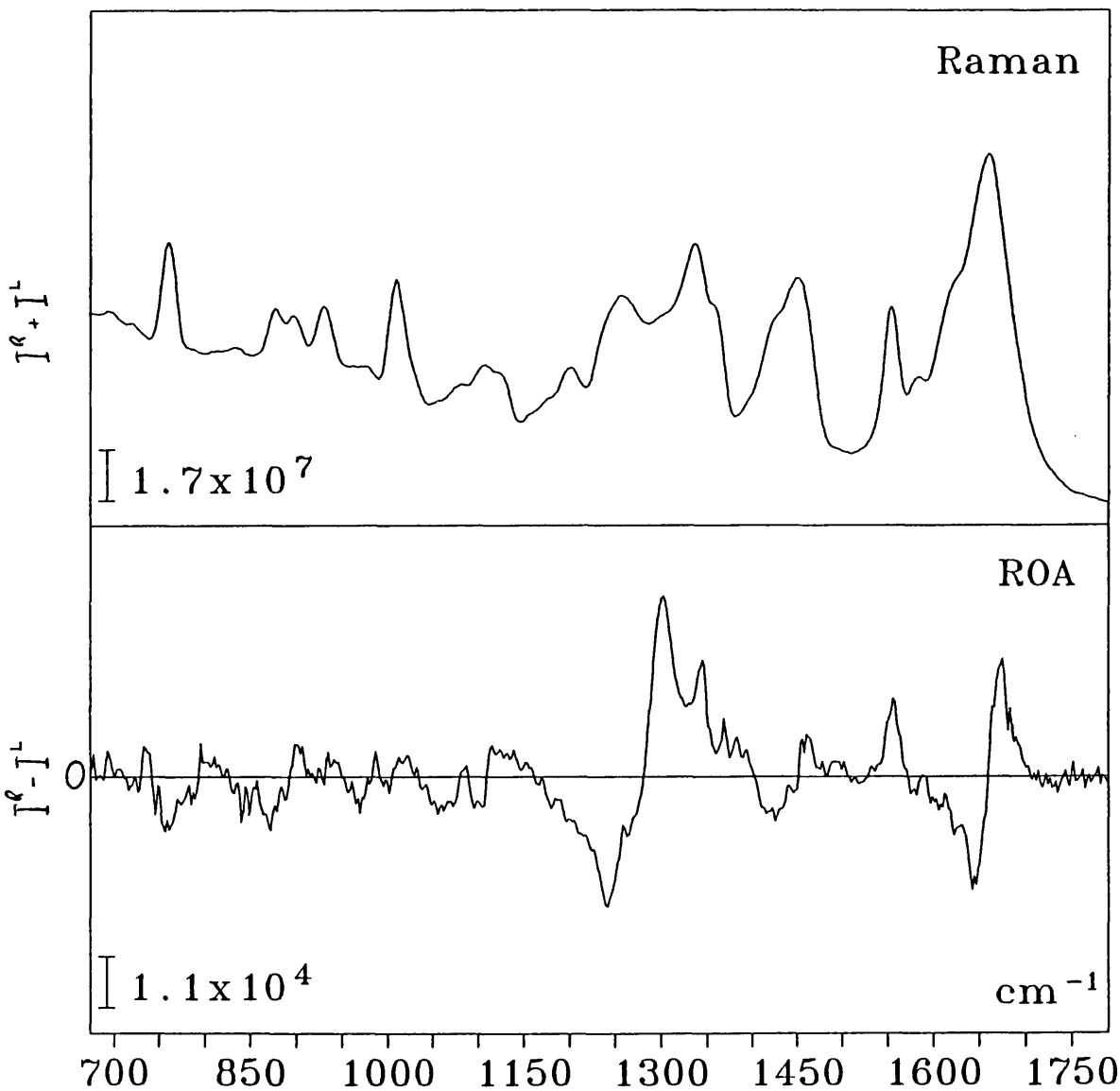


Figure 6.5: The Raman and ROA spectra of HEWL.

tryptophan benzene ring stretch and pyrrole C^2-C^3 stretch modes, respectively (Harada and Takeuchi, 1986).

The amide I ROA is a negative/positive couplet, between 1640 cm^{-1} and 1700 cm^{-1}. Although a couplet of this kind is seen in all protein ROA the relative intensities of the negative and positive regions of the couplet vary and it has been proposed (Wen et al., 1994a) that this relates to the secondary structure content of a protein. (This relationship is further discussed in chapter 8.) The amide I ROA alters significantly upon denaturation, and slightly upon deuteration.

At 1620 cm^{-1} there is a negative ROA band that forms a shoulder to the negative component of the amide I band. This band may originate from the amide I or the aromatic vibrations. The latter is the most likely because the band appears to stay at the same frequency upon deuteration and amide I vibrations do not occur so low.

The positive/negative ROA couplet at 1550 to 1580 cm^{-1} corresponds to the two tryptophan bands at 1555 and 1580 cm^{-1}. The former band has quite a significant ROA signal, probably because this vibration (C^2-C^3 stretch) is close to the chiral C_α.

1500-1150 cm^{-1}

The first band, at 1451 cm^{-1}, in this region is the CH_2 deformation band (Lord and Yu, 1970) which is large in all protein samples and is analogous to the CH_2 and CH_3 deformation bands of the peptide samples. There is a low frequency shoulder on this band, at 1430 cm^{-1}, which has been assigned to a tryptophan mode (Lord and Yu, 1970) involving the N^1-C^2-C^3 symmetric stretch, NH bending and aromatic C-H bending (Harada and Takeuchi, 1986).

The ROA in this region shows a negative/positive couplet between 1400 and 1475 cm^{-1} which comes from the CH_2 vibrations. Althought the negative component of this couplet could be the 1430 cm^{-1} tryptophan vibration it is an unlikely assignment because the couplet is still present when this tryptophan signal is removed by deuteration and a negative signal is also found in this region in BSA and insulin (which have 2 tryptophans and no tryptophans respectively: G. Wilson, unpublished data). Although this band has always been seen as a couplet, the intensity of the two sides has varied as the HEWL has been run in different instrumental set ups.

The next large Raman band occurs at 1340 cm^{-1}, and it has associated with it a high frequency shoulder at 1360 cm^{-1}. These bands are a Fermi resonance doublet from the tryptophan residues, arising from the fundamental vibration of the N^1-C^8 stretching mode and combinations of low frequency C-H bending modes (Harada and Takeuchi, 1986). The ratio of the peak heights is an indication of the tryptophan environment (Harada et al., 1986). However, because the low frequency peak is superimposed upon the high frequency amide III/C-H bend vibrations, it is not possible to use the ratio as an indicator of tryptophan environment in proteins. The 1340 cm^{-1} band has a small low frequency shoulder at ~1300 cm^{-1}. Since this band cannot be found in the deuterated Raman spectra where the low frequency side of the tryptophan band is clear, it is probably one of the amide III vibrations.

The broad non-symmetrical Raman band at 1260 cm^{-1} is the amide III band. This assignment is easily proven by its disappearance upon deuteration and its absence in the summation Raman spectra of the constituent amino acids of HEWL (Lord and Yu, 1970).

The amide III ROA consists of a broad negative/positive couplet between 1200 and 1400 cm^{-1}. This couplet is found in all proteins, but the fine structure superimposed on the couplet is different. The amide III region can be classed as the key region for protein ROA. The amide III ROA is more reliable than the amide amide I ROA because the amide III vibrations are non polarized. The amide III ROA is one of the more reproducible regions of the spectrum, and is usually the first to appear in the protein ROA acquisition. However, complete assignments are not given in this section because they are too complicated; instead they are discussed later in chapter 8.

In HEWL positive ROA peaks are seen at 1340 and 1300 cm^{-1}. There are also small peaks which appear at a higher frequency, between 1360 and 1400 cm^{-1} and although small, they consistently appear in HEWL spectra. The crossover point of the broad couplet is at approx. 1280 cm^{-1}. The negative couplet shows a small, but again consistent, negative peak at 1270 cm^{-1}. The largest negative peak appears at 1245 cm^{-1} and to the low frequency side of this is a broad positive/negative couplet between 1110 and 1210 cm^{-1}.

There still remains the question whether the ROA signals we see in this region originate from amide links, i.e. amide III ROA, or whether they come from sidechain vibrations, such as tryptophan. Ideally, there should be very few ROA signals from sidechain vibrations, but they do occur in other regions of the spectrum and so we should examine the possiblity of their appearance in this region. The sidechain vibration that occurs in the Raman amide III region is the Fermi resonance of tryptophan at 1340/1360 cm^{-1} (particularly worrying because tryptophan has ROA active signals in other regions of the spectrum). There are three pieces of evidence that point to the 1340 cm^{-1} ROA band being a genuine amide III signal; first, a

1340 cm^{-1} ROA signal appears in BSA which has only 2 tryptophans in 582 amino acids and shows no tryptophan Raman or ROA in the 1550 to 1570 cm^{-1} region; second, the deuterium exchange experiments show that the when all the tryptophans should be exchanged the 1340 cm^{-1} signal can still be seen (see section 6.3.5), but when the protein is completely exchanged this signal vanishes; third, upon denaturation the tryptophan ROA at 1550 to 1570 cm^{-1} remains whereas the 1340 cm^{-1} band vanishes.

The next Raman band occurs at 1200 cm^{-1}, with low frequency shoulders at 1175 cm^{-1} and 1160 cm^{-1}. The 1200 cm^{-1} band can be assigned to the ring-C stretches of tyrosine (1210 cm^{-1}) and phenylalanine (1207 cm^{-1}) (Lord and Yu, 1970; Harada and Takeuchi, 1986). The low frequency bands may be in plane C-H bends of the aromatic residues (1180 cm^{-1}, 1183 cm^{-1} and 1148 cm^{-1} in tyrosine, phenylalanine and tryptophan, respectively: Harada and Takeuchi, 1986), but they are not observed in the summation spectra of Lord and Yu (1970). Another possiblity is that they are C-N stretch backbone vibrations similar to those that occur in this region in the peptides.

1150-700 cm^{-1}

The next three Raman bands are at 1125 cm^{-1}, 1110 cm^{-1} and 1075 cm^{-1}. Only the central band is affected by deuteration and denaturation. The summation Raman spectra of Lord and Yu (1970) shows that the first and last bands have sidechain residue origins (tryptophan and tyrosine C-H bends for the 1075 cm^{-1} band and methyl rocks from Val, Ile etc. for the 1125 cm^{-1} band). However, Lord assigns all these bands to C-N stretches. In the peptides this region showed sensitivity to deuteration and it was proposed that bands in this region were due to methyl and NH_3^+ rock and C-N stretching. Therefore, it is likely that in proteins these vibrations

come from a similar combination of sidechain vibrations and C-N stretching.

The ROA of this region consists of a small negative/positive couplet centred at 1100 cm^{-1} giving positive ROA at 1125 cm^{-1}, and a broad negative band at about 1050 cm^{-1}. The behaviour of these bands upon deuteration and denaturation mirrors the behaviour of the Raman bands of the same frequency: the 1100 cm^{-1} couplet is related to the 1110 cm^{-1} Raman band, and the negative 1050 cm^{-1} signal is related to the 1075 cm^{-1} Raman band. This indicates that the Raman and ROA bands are probably quite closely related. The 1100 cm^{-1} couplet comes mainly from backbone C-N stretch with N-H involvement. This would explain the alteration of this band with denaturation and deuteration. The 1125 and 1075 cm^{-1} Raman and ROA features come either from sidechain modes, or deuterium exchange and conformationally insensitive backbone vibrations (although this is less likely considering that the peptide backbone vibrations show great sensitivity to deuterium exchange in this region).

The Raman band at 1015 cm^{-1} is assigned to ring breathing vibrations of the tryptophan and phenylalanine benzene rings. It is slightly asymmetic because of a high frequency aromatic C-H bend from phenylalanine at 1032 cm^{-1} (Harada and Takeuchi, 1986). The band at 930 and 900 cm^{-1} are found in HEWL spectra and assigned to C-C stretching. However, in this sample these bands are exaggerated due to underlying similar bands from the acetate buffers. The 880 cm^{-1} band originates from benzene ring and N-H vibrations of tryptophan.

The 1015 cm^{-1} tryptophan and phenylalanine vibration has a weak, high frequency, positive ROA signal associated with it at 1025 cm^{-1}. In insulin a positive ROA feature is associated with the high frequency phenylalanine vibration. The small negative ROA feature at 1015 cm^{-1} maybe genuine.

Insulin and BSA both have quite small features in this region (Wen et al., 1994a) and this is because these vibrations originate far from the chiral C_α. The 930 and 900 cm^{-1} Raman bands originate in C-C stretches and they both have positive ROA associated with them. This is confirmed by the ROA spectrum of HEWL in H_2O where there are no acetate bands to interfere (Barron et al., 1992a). Both insulin and BSA (G.Wilson, unpublished data), which have high alpha helical content, shown large positive ROA in this region, and a survey of other proteins (Barron et al., 1992, Wen et al., 1994b), bear out the proposed relationship (Wen et al., 1994a) between positive ROA in this region and α-helix content. However, great care must be exercised with these assignments because baseline redrawing procedures may introduce artifacts. These Raman bands have been related to α-helical content (Tu, 1986; Heremans and Wong, 1985).

The 880 cm^{-1} tryptophan band shows negative ROA. Neither BSA or insulin (G.Wilson, unpublished data) show an ROA feature in this region.

The 830 cm^{-1} bands is a Fermi resonance doublet from tyrosine. The fundamental band is a ring symmetric stretch coupled to C-C and C-O stretch vibrations. The overtone is a non-planer ring vibration which originates at 413 cm^{-1} (Siamwiza et al., 1975). This low frequency vibration has a large contibution from the hydroxyl group which makes this Fermi resonace doublet sensitive to the hydrogen bonding of the environment.

The 765 cm^{-1} vibration is the indole ring breathing of tryptophan and gives a negative ROA band. A small band appears at 715 cm^{-1} which disappears upon deuteration. In the peptides deuterium sensitive bands in this region are assigned to CO_2^- vibrations. The 690 cm^{-1} band is the C-S stretch from methionine (Lord and Yu, 1970).

6.3.2 Deuterated HEWL

The Raman and ROA spectra of deuterated HEWL is shown in Figure 6.6.

1750-1550 cm^{-1}

The amide I' band is at 1655 cm^{-1}, but a high frequency shoulder appears in the deuterated spectrum at 1710 cm^{-1}. This latter band is CO_2D carbonyl stretch probably from both the aspartate and glutamate residues in HEWL and the acetate buffer (see the Ala_2 results). The amide I' band is 10 cm^{-1} lower than the amide I band and as the amide vibration moves down so it obscures the aromatic residue ring stretch, with can only just be seen in the deuterated HEWL. The amide I' frequency is lower than the amide I because of the small involvement of N-H in plane bend in the latter. The amide I ROA also moves down by 10 cm^{-1} and this movement obscures any 1620 cm^{-1} aromatic ROA that may be present. The amide I' ROA is more intense than it's hydrated counterpart, but the reason for this is unclear.

The remaining two tryptophan vibrations occur at 1570 cm^{-1} and 1555 cm^{-1}. The high frequency band drops slightly upon deuteration. Such a shift in this band is expected (Harada and Takeuchi, 1986). The frequency of the equivalent ROA vibrations follows a similar pattern. The intensity of the couplet increases, as does the intensity of all the ROA in the region 1350 to 1700 cm^{-1}, but again the reason for this intensity increase is not known.

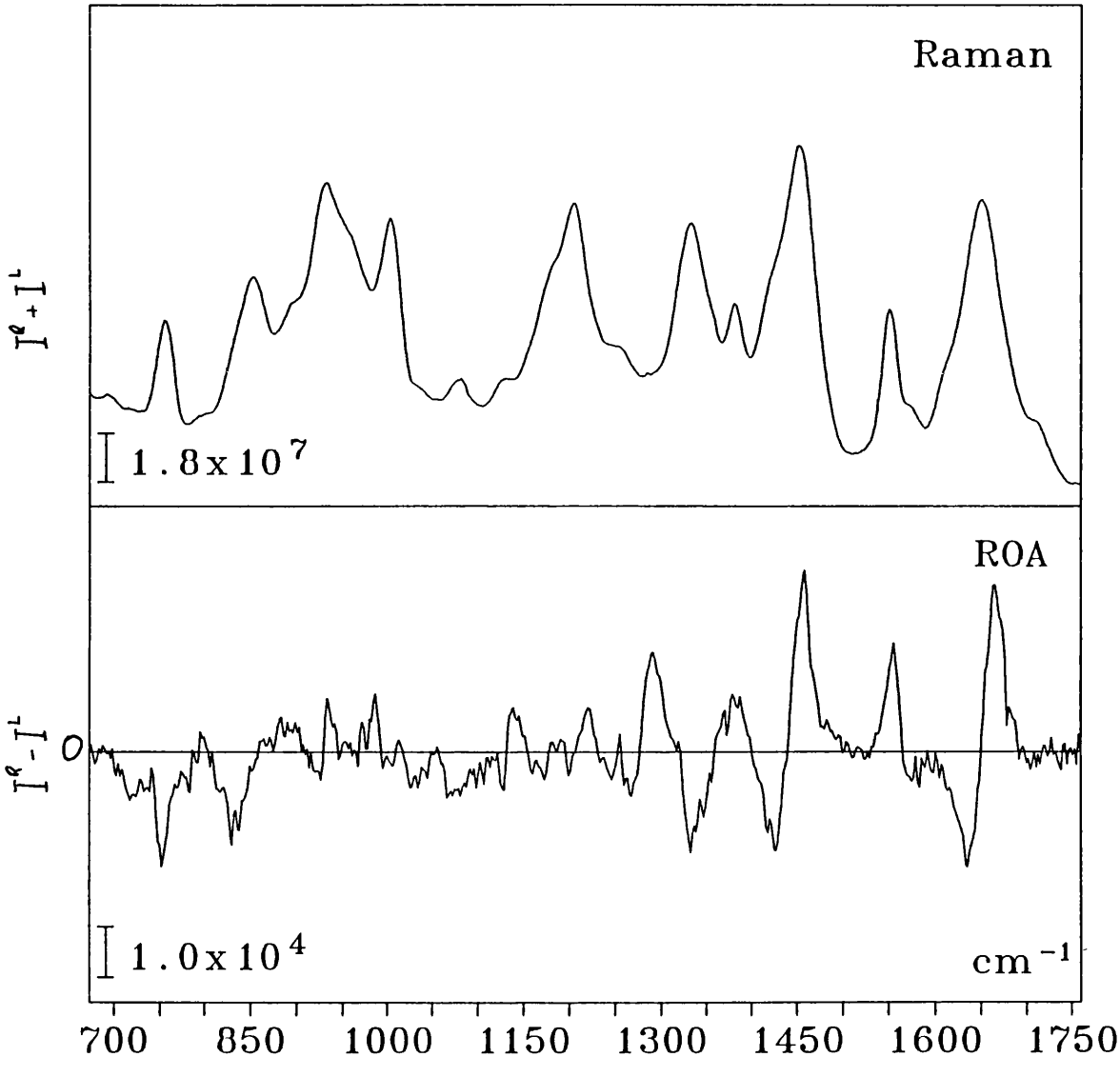


Figure 6.6: The Raman and ROA spectra of completely deuterated HEWL.

1500-1150 cm^{-1}

The 1451 cm^{-1} CH_2 deformation band remains at the same frequency upon deuteration. However, it increases in intensity due to mixing with the amide II' vibration (as occurred in the peptides). The peak is slightly asymmetric but this cannot be tryptophan vibrations because upon deuteration the tryptophan band, at 1430 cm^{-1}, falls in frequency by approx 50 cm^{-1} to 1380 cm^{-1} (Harada and Takeuchi, 1986). This asymmetry was observed by Lord and Yu in 1970 and assigned to different C-H deformations.

The deuteration of HEWL leads to an enhancement of the ROA couplet (a similar process occurs with the peptides) which, as with the Raman enhancement, is caused by the amide II' band. The magnitude of the ROA enhancement is similar to the Raman enhancement with the CID value for this band rising by only 50% upon deuteration.

In the deuterated HEWL a tryptophan vibration appears at 1375 cm^{-1}. This band is characteristic of deuterated tryptophan, and is due to the N^1-C^2-C^3 symmetric stretch. Upon deuteration the positive ROA peaks between 1350 and 1400 cm^{-1} in the hydrated HEWL coalesce and become slightly more intense. When the partially deuterated samples are examined it become apparent that this effect depends on the extent of deuteration. Consideration of the tryptophan exchange rates (Wedin et al., 1982) indicates that this ROA peak originates from the amide link exchange effect rather than from the tryptophan exchange. However, the ROA intensity between 1350 and 1400 cm^{-1} is small and the above argument would be invalidated if small artifacts had increased, or decreased, intensity in this region with any of the deuterated samples (perhaps by baseline correction). The tryptophan vibrations themselves become more

localized around the N^1-C^2-C^3 bond upon deuteration and this may give rise to an ROA signal where there was previously none.

The tryptophan Fermi resonance doublet also changes upon deuteration, becoming a single, asymmetric, peak at 1335 cm^{-1}. In solvated tryptophan the doublet drops by ~8 cm^{-1} upon deuteration of N^1 but with tryptophan derivatives the behaviour is less obvious (Harada et al., 1986).

The deuteration of the amide links will alter the vibrations in the amide III region, and as the N-H deformation moves down to 900 cm^{-1}, it will leave behind uncoupled C-H deformations. The remaining C-H deformations may underlie most of this region (the peptide studies show that C-H deformations can range from 1355 to 1245 cm^{-1}). In HEWL the C-H deformations appear at 1205 cm^{-1}. The high frequency shoulder at 1250 cm^{-1} may be other C-H vibrations, or a band previously obscured by the amide III band. The low frequency Raman band 1175 cm^{-1} is probably the ring-C stretches from tyrosine and phenylalanine, or aromatic C-H deformations.

The low frequency negative component and the high frequency positive component of the broad amide III ROA couplet collapse upon deuteration to give a small positive peak and a larger positive/negative couplet, respectively. Similar changes are seen in the peptides upon deuteration. The positive 1340 cm^{-1} band in hydrated HEWL becomes a large negative band at the same frequency. The positive 1300 cm^{-1} band falls in intensity and frequency to about 1290 cm^{-1}. The negative 1270 cm^{-1} band remains, rising by 5 cm^{-1} upon complete deuteration. In the deuterated HEWL the negative 1245 cm^{-1} band of hydrated HEWL becomes a positive 1220 cm^{-1} band. The small negative and positive features that are seen between 1100 and 1275 cm are reproducable and are seen in previous experiments.

The low frequency positive component of the hydrated HEWL 1110 to 1210 cm^{-1} couplet remains positive upon deuteration, but does move to a higher frequency. The high frequency negative component collapses to form a small, but reproducible, positive ROA peak.

1150-700 cm^{-1}

The first two bands in this region are 1130 cm^{-1} and 1075 cm^{-1}. The first band has increased in frequency by 5 cm^{-1} upon deuteration and the latter has remained invariant. This indicates mixing between the sidechain methyl group and the C-N stretching, while the C-H bends of the aromatic groups remain more isolated. Upon deuteration the 1110 cm^{-1} band of the hydrated HEWL has disappeared both in the Raman and ROA spectrum. This would indicate that this band originates from deuterium sensitive backbone modes; the sidechain vibrations in this region come from the non-exchangable aromatic parts of the residues. The 1130 cm^{-1} Raman band seems to maintain its postive ROA signal (although it is possible that there is a slight S-artifact at 1130 cm^{-1}). The 1075 cm^{-1} band maintains it's negative ROA signal.

The changes in ROA in this overall region are slightly speculative because in the partially deuterated samples there appears to be a breakdown in the consistency of the changes that occur as deuteration progresses (see section 6.3.4).

The tryptophan and phenylalanine band at 1015 cm remains the same upon deuteration. The ROA associated with this band is unclear.

The 930 cm^{-1} band has a high and low frequency shoulder at 960 cm^{-1} and 900 cm^{-1}, respectively. The 930 cm^{-1} and 900 cm^{-1} bands are C-C

stretch, but these bands overlie broad N-D in-plane deformations which are responsible for the 960 cm^{-1} shoulder (Lord and Yu, 1970).

The 850 cm^{-1} band is a mixed benzene and N-D vibration (Harada and Takeuchi, 1986), which was observed at 880 cm^{-1} in the hydrated HEWL. This band appears to have gained in intensity, probably because it is sitting on top of broader N-D in plane bends. However, this gain is not observed by Miura and co-workers (Miura et al., 1991). (It is not immediately clear from the experimental procedures described by Muria et al. what Raman experiment they were doing, i.e. scattering angle, polarization filters used, etc.) The ROA of this band is a negative peak which also moves down upon deuteration.

The 765 cm^{-1} tryptophan band remains invariant upon deuteration and still gives negative ROA. The 715 cm^{-1} band almost disappears, but the methionine C-S stretch at 690 cm^{-1} remains.

6.3.3 Denatured HEWL

The Raman and ROA spectra of denatured HEWL are shown in Figures 6.7 to 6.9.

The study of the denaturation of HEWL by Raman is not new. In 1973 and 1974 Chen and co-workers (Chen et al., 1973 and 1974) published the Raman spectra of HEWL which had been irreversibly thermally denatured in the former and chemically denatured in the latter. However, there is one point to make about the terminology used in this former paper. In it Chen and co-workers claim (citing ORD and NMR papers) that HEWL is 'reversibly' denatured between 32°C and 76°C and 'irreversibly' denatured above 80°C. This, I believe, is erroneous; differential scanning calorimetry data shows that at pH 5 HEWL is denatured at 77°C (Cooper et al, 1992a).

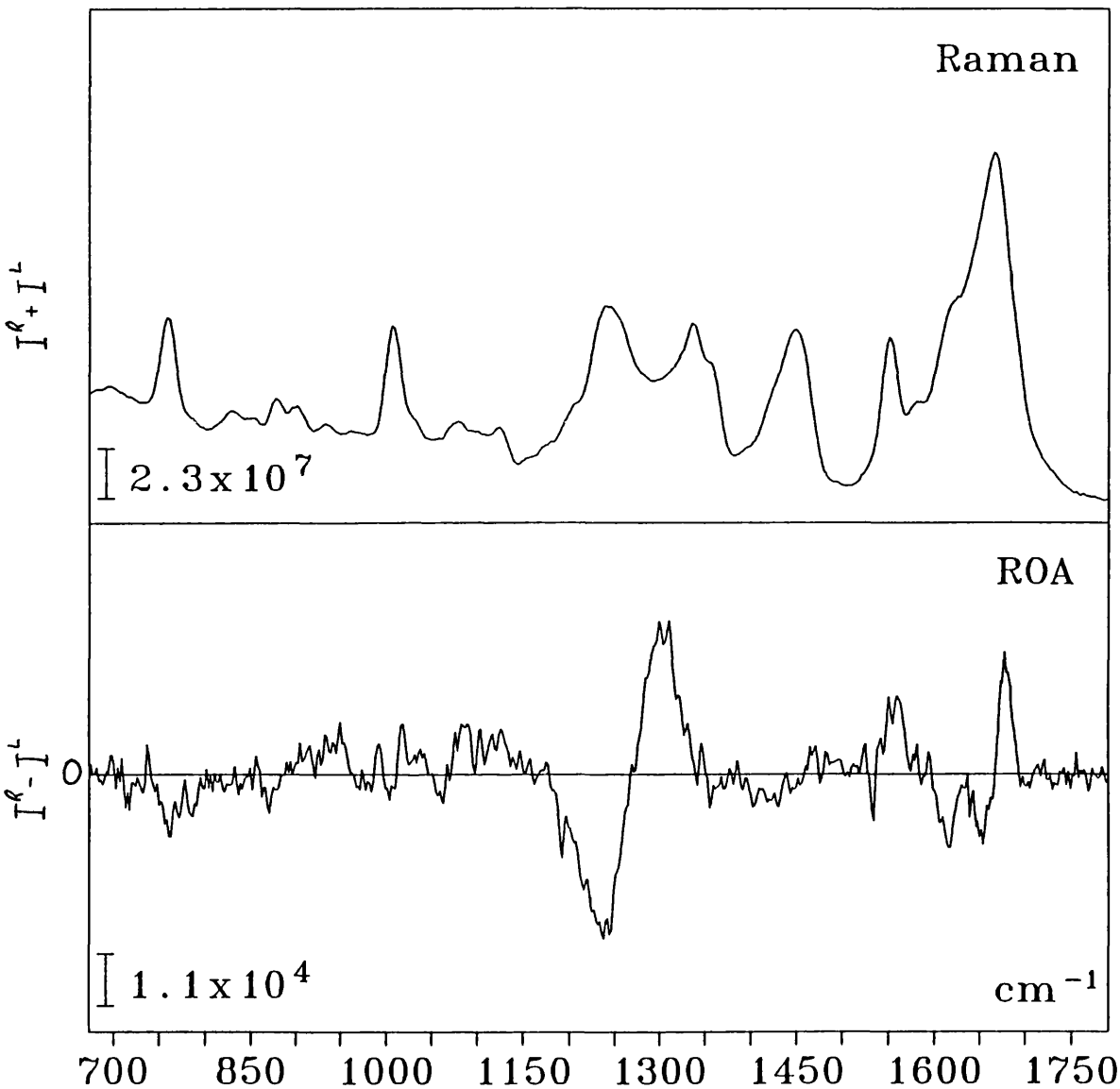


Figure 6.7: The Raman and ROA spectra of denatured HEWL.

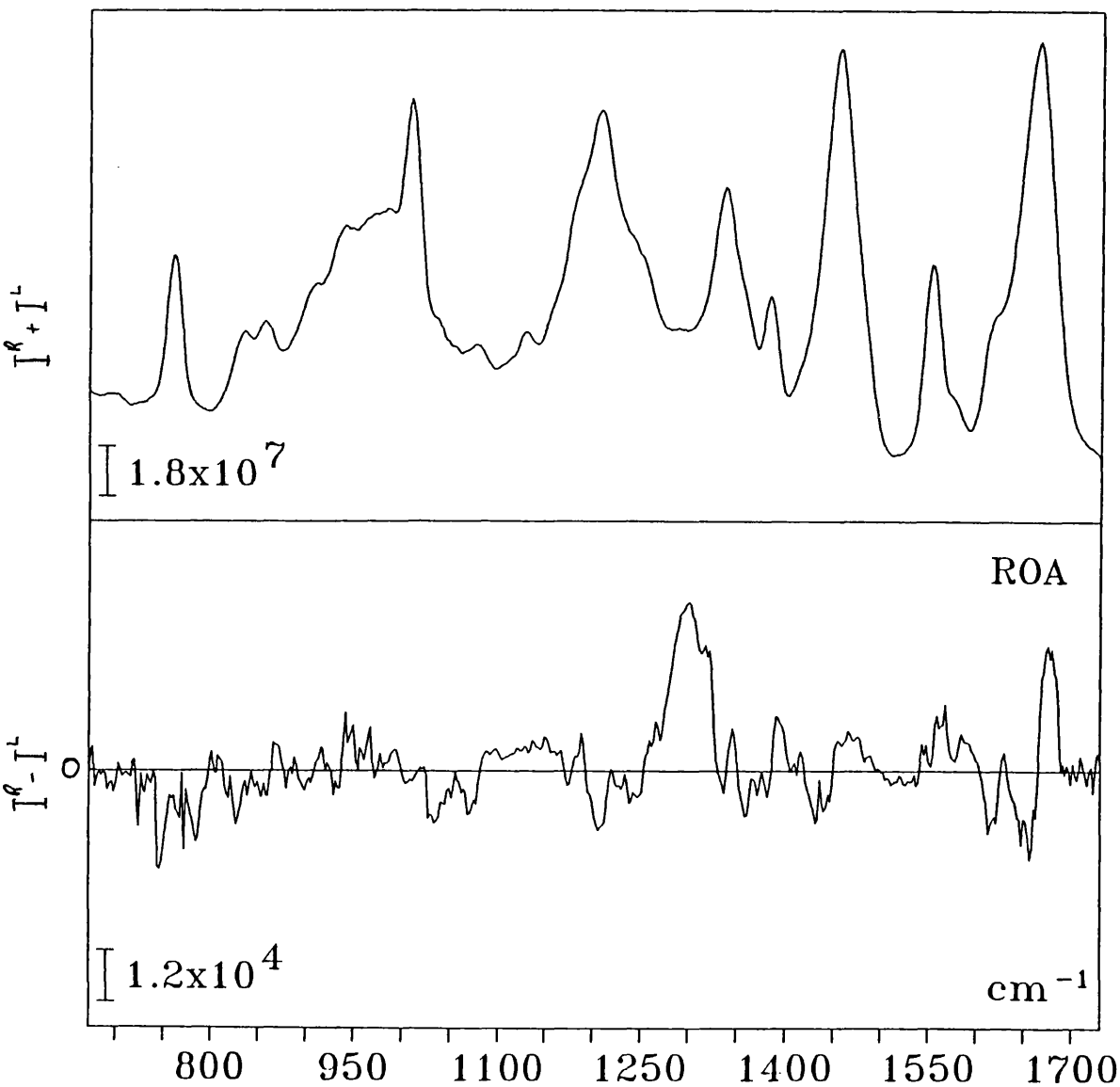


Figure 6.8: The Raman and ROA spectra of completely deuterated, denatured HEWL. (I would like to acknowledge the assistance of G. Wilson in obtaining the above spectra.)

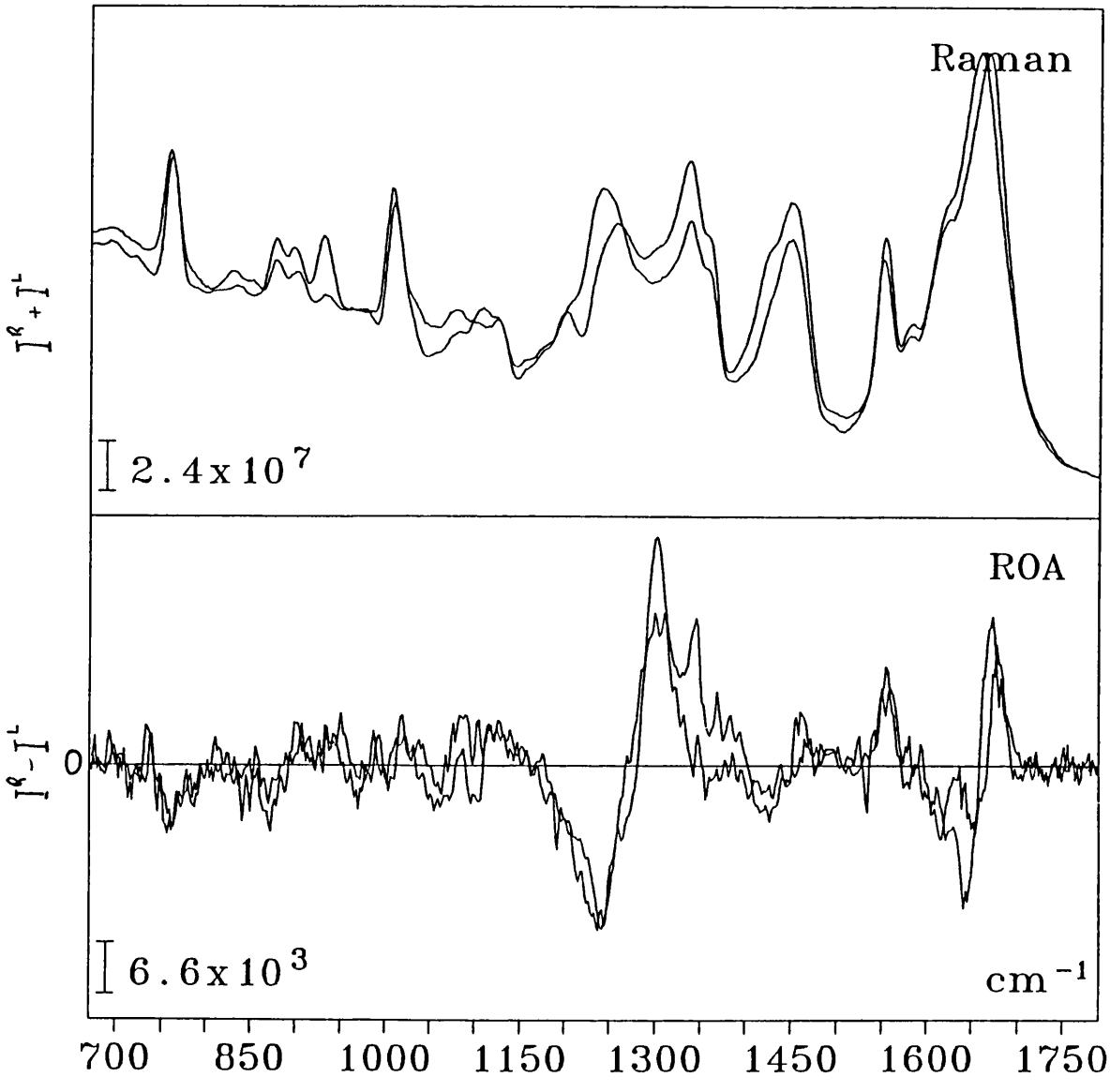


Figure 6.9: The comparison of the Raman and ROA spectra for HEWL and denatured HEWL.

It follows that Chen's 'reversibly' denaturation was not denaturation at all, and that true denaturation only occurred above 80°C where it was 'irreversible' because the high concentration caused gelation. Their Raman results bear this out, because they see no alteration in the Raman spectra until they 'irreversibly' denature their HEWL. In fact they conclude themselves that;

"reversible denaturation of aqueous lysozyme at temperatures up to 76°C for periods of time up to 2-3 hrs produces hardly any conformational change that is detectable by Raman spectoscopy".

1750-1500 cm^{-1}

Upon denaturation the amide I band moves up from 1665 cm^{-1} to 1670 cm^{-1}. This was observed by Chen and co-workers. This increases in frequency is caused by changes in the hydrogen bonding network of the protein. It is known that when the hydrogen bonding to the amide link gets weaker the amide I frequency rises (Tu, 1986). As the amide I band moves to higher frequency, so the low frequency shoulder from the aromatic residues becomes more prominent because the aromatic residue ring stretch remains at the same frequency as it did in the native HEWL, as do the two lower tryptophan bands. The constant frequency of the lowest tryptophan band at 1555 cm^{-1} may indicate that the sidechain conformation of the tryptophans remains similar between the native and denatured states. This band has been shown to be conformationally sensitive and the frequency can vary by as much as 15 cm^{-1} (Miura et al., 1989).

The amide I ROA couplet also increases in frequency by about 10 cm^{-1}. Although the couplet decreases in intensity upon denaturation, the negative component decreases more drastically, becoming half the size of the

positive component. The aromatic ring stretch ROA remains at the same intensity but moves down slightly by about 5 cm^{-1}. This may represent a change in the conformation or environment of the aromatic residues, or be caused by the changes in the adjacent amide I couplet.

It is appropriate here to consider the ROA amide I/protein structure relationships that the native and denatured HEWL spectra implies. Overall, the changes in the amide I ROA indicate that it is sensitive to protein structure, with the denatured protein producing this type of small negative/large positive couplet. However, a similar decrease in the negative intensity is found in the molten globule form of α-lactalbumin, which some residual structure (see chapter 7). It seems unlikely that the denatured form of HEWL has any residual secondary structure (see 6.1.6) but UVCD (another chiroptical technique) also shows secondary structural elements in reduced HEWL (White, 1982). In the denatured state the individual amide links will adopt conformations determined by the minima in the energy landscape. So although extended secondary structure maybe absent it does not preclude the selection of Φ and Ψ angles that represent α-helix and β-sheet by individual amides. Another possibility is the formation of transient secondary structural elements. Unfortunately, the conclusions drawn out here are not completely definitive because the amide I ROA is experimentally difficult to obtain accurately. The amide I ROA is further discussed in chapter 8.

The tryptophan ROA couplet between 1550 and 1580 cm^{-1} loses its negative high frequency component. This may be an experimental artifact, but could also represent a genuine change in tryptophan structure or environment.

1500-1150 cm^{-1}

The 1451 cm^{-1} CH_2 deformation band remains at the same frequency in the denatured state as in the native state. The adjacent tryptophan band, which was at 1430 cm^{-1} in the native state, is less prominent in the denatured state. Whether this is because of an increase in frequency, or a decrease in intensity is unclear from our Raman spectra. Earlier work on reduced HEWL powder (Chen et al, 1974) indicates no change. A similar change happens when bovine α-lactalbumin (BAL) is converted into it's molten globule state. This tryptophan band is sensitive to hydrogen bonding (Miura et al., 1989).

The ROA couplet that represents the CH_2 and CH_3 deformations drops in intensity, but maintains a similar frequency. This decrease represents the averaging of the sidechain conformations.

The 1340/1360 cm^{-1} Fermi resonance doublet appears at the same frequency as in the native state, but drops slightly in intensity. This was observed by Chen et al. (1973) and it was proposed that this reflected the increase in solvent exposure of the tryptophans. This conclusion is in agreement with the BAL studies (Yu, 1974) and tryptophan work (Miura et al., 1989). The underlying amide III vibrations complicate further analysis because they will also change upon denaturation.

The next large Raman band occurs at 1240 cm^{-1} and is easily assigned to the amide III vibration of a random coil which occurs in this region (Tu, 1986) and has been observed in earlier work (Chen et al., 1973 and 1974). This decrease is because the hydrogen bonding to the amide links becomes weaker and this lowers the N-H bending frequency. Interestingly, a Raman study of pressure induced denaturation of HEWL showed the

opposite of this (Heremans and Wong, 1985) with a decrease in the 1245 cm^{-1} band while the 1335 cm^{-1} band increased.

The denatured amide III vibration shows a broad negative/positive ROA couplet. The negative peak at 1245 cm^{-1} remains at the same intensity as the native HEWL, the crossover is at 1270 cm^{-1} and the intensity of the positive peak at about 1305 cm^{-1} drops by about 30% from the native form. The absence of the positive 1340 cm^{-1} band is of great interest. The 1270 cm^{-1} negative band in the native state vanishes upon denaturation. The assignments of these amide III ROA bands are discussed at the end of this chapter. The broad amide III couplet represents a superposition of 'snapshot' ROA spectra of a protein with large dynamic motions. The deuterated denatured HEWL ROA spectra show similar features to that of the native HEWL, with the disappearance of the negative component of this couplet.

The amide III band has low frequency shoulders at 1210 cm^{-1}, 1175 cm^{-1} and 1160 cm^{-1}. The assignment of these bands is the same as that for the native state of the protein.

1150-700 cm^{-1}

The next three bands are at 1125 cm^{-1}, 1095 cm^{-1} (15 cm^{-1} lower than in the native state) and 1075 cm^{-1}. The high and low frequency bands remain at a similar intensity, while the central band collapses. The ROA features for the high and low frequency bands remain, while the ROA of the middle band collapses. This reinforces the assignment of this central band to a conformationally (and deuteration) sensitive backbone vibration.

The phenylalanine and tryptophan band at 1005 cm^{-1} is slightly lower than in the native state. This is due to the tryptophan vibration moving

down and merging with the phenylalanine vibration (Chen et al., 1973). This reveals more of the the 1030 cm^{-1} phenylalanine band. The ROA feature appears to remain similar between the native and denatured state.

The following two C-C stretch bands at 930 cm^{-1} and 900 cm^{-1} remain at the same frequency as in the native state. This is complicated by the interference of the buffer bands at this frequency. The ROA of the 930 and 900 cm^{-1} bands is positive. This questions the hypothesis that positive ROA in this region represents α-helix, but HEWL is a poor example of an α-helical protein and the positive region may be artifically induced by baseline correction.

The 880 cm^{-1} tryptophan Raman band decreases in intensity as expected for denatured proteins (Chen et al., 1973 and 1974), but the frequency remains the same. However, the frequency of this band is believed to be sensitive to environment (Harada and Takeuchi, 1986) and and so this indicates either an error in the tryptophan analysis or a small change in the tryptophan environments upon deuteration. (NMR results on thermally dentaured HEWL show that the tryptophans are still partially buried, but have a substantial amount of conformational averaging (Radford et al., 1992; Evans et al., 1991).)

Two bands appear at 850 cm^{-1} and 825 cm^{-1}, from the single broad 830 cm^{-1} tyrosine Fermi resonance band, upon denaturation. This was observed by Chen and co-workers in thermally denatured HEWL, but at the time this tyrosine band had not been studied in depth. Later studies on tyrosine (Siamwiza et al., 1975) showed that the doublet was sensitive to hydrogen bonding, but not to environment or tyrosine conformation. Therefore, the changes that we see are due to hydrogen bonding with the solvent. (Chen's spectra, plus Siamwiza's results, indicate that upon thermal denaturation tyrosines are become hydrogen bond acceptors.)

The 760 cm^{-1} band is the indole ring breathing vibration of tryptophan. Both its intensity and its ROA signal remain the same upon denaturation. A slight hint of the 715 cm^{-1} band remains in the denatured state. The 690 cm^{-1} band is the C-S stretch from methionine.

6.3.4 HEWL-NAG_3 experiments.

The Raman and ROA spectra for HEWL-NAG_3 are shown in Figures 6.10 and 6.11. The HEWL inhibitor experiment has been run three times and the results of all three of these experiments are very similar. The Raman and ROA spectra of NAG_3 have also been obtained, and these are shown together with a (HEWL-NAG_3)-NAG_3 subtraction spectra in Figures 6.12 and 6.13.

1750-1500 cm^{-1}

The Raman and ROA spectra of the HEWL and HEWL-NAG_3 are similar. There appear to be no differences between the two samples in the region 1750 to 1500 cm^{-1}. There may be some changes in the aromatic ROA between 1570 and 1630 cm^{-1} but these changes have were not seen in earlier experiments.

1500-1150 cm^{-1}

In the Raman spectra of HEWL-NAG_3 the low frequency tryptophan and CH_2 deformation shoulder at 1430 cm^{-1} appears to be more enhanced. Additionally, more Raman intensity appears between this band and the 1360 cm^{-1} tryptophan vibration. The subtraction spectra show that both these effects are due to the NAG_3 vibrations alone.

It is in the amide III region that some of the most interesting changes occur. In the Raman spectrum the 1360 cm^{-1} tryptophan band becomes

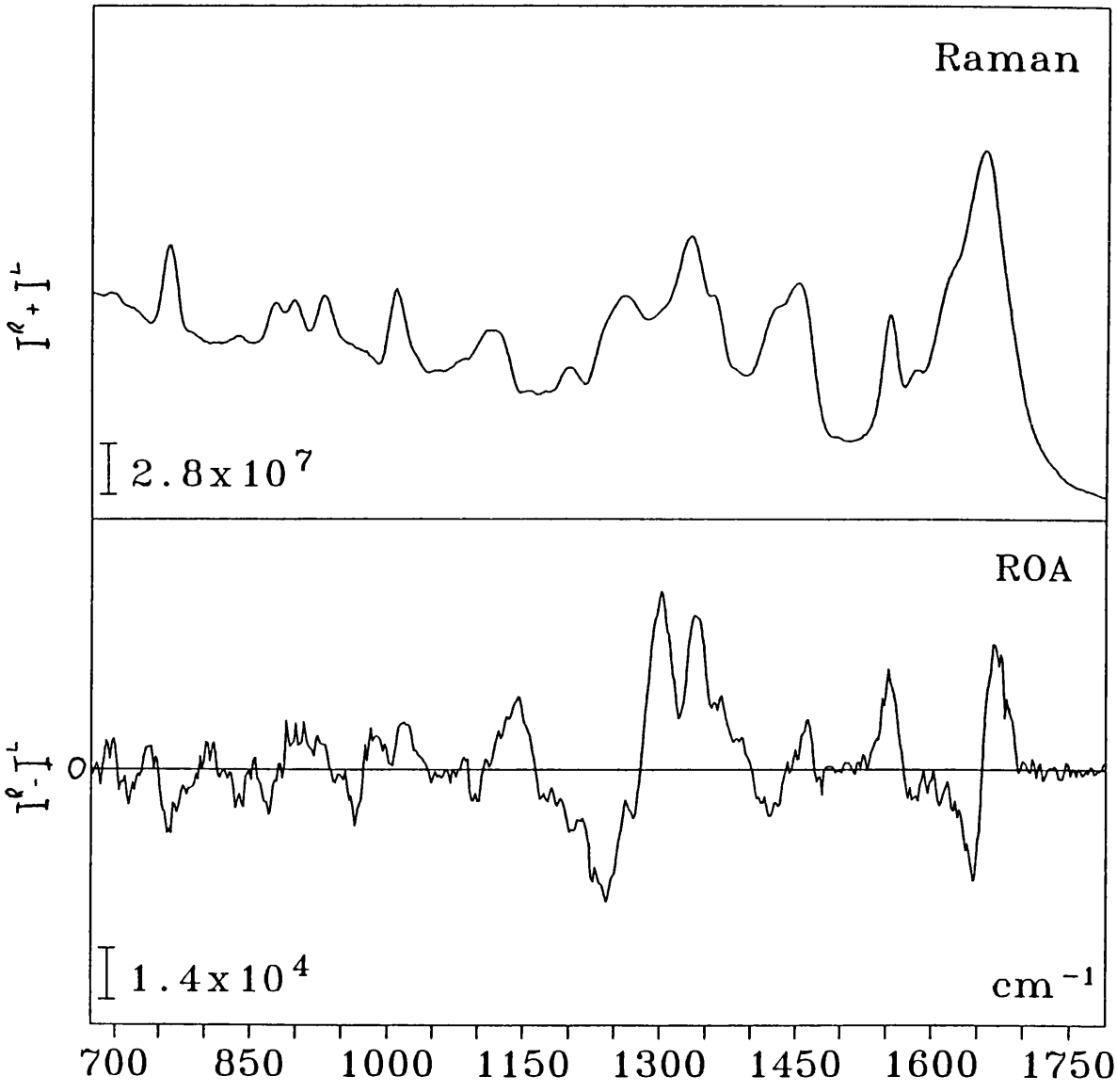


Figure 6.10: The Raman and ROA spectra of HEWL-NAG_3.

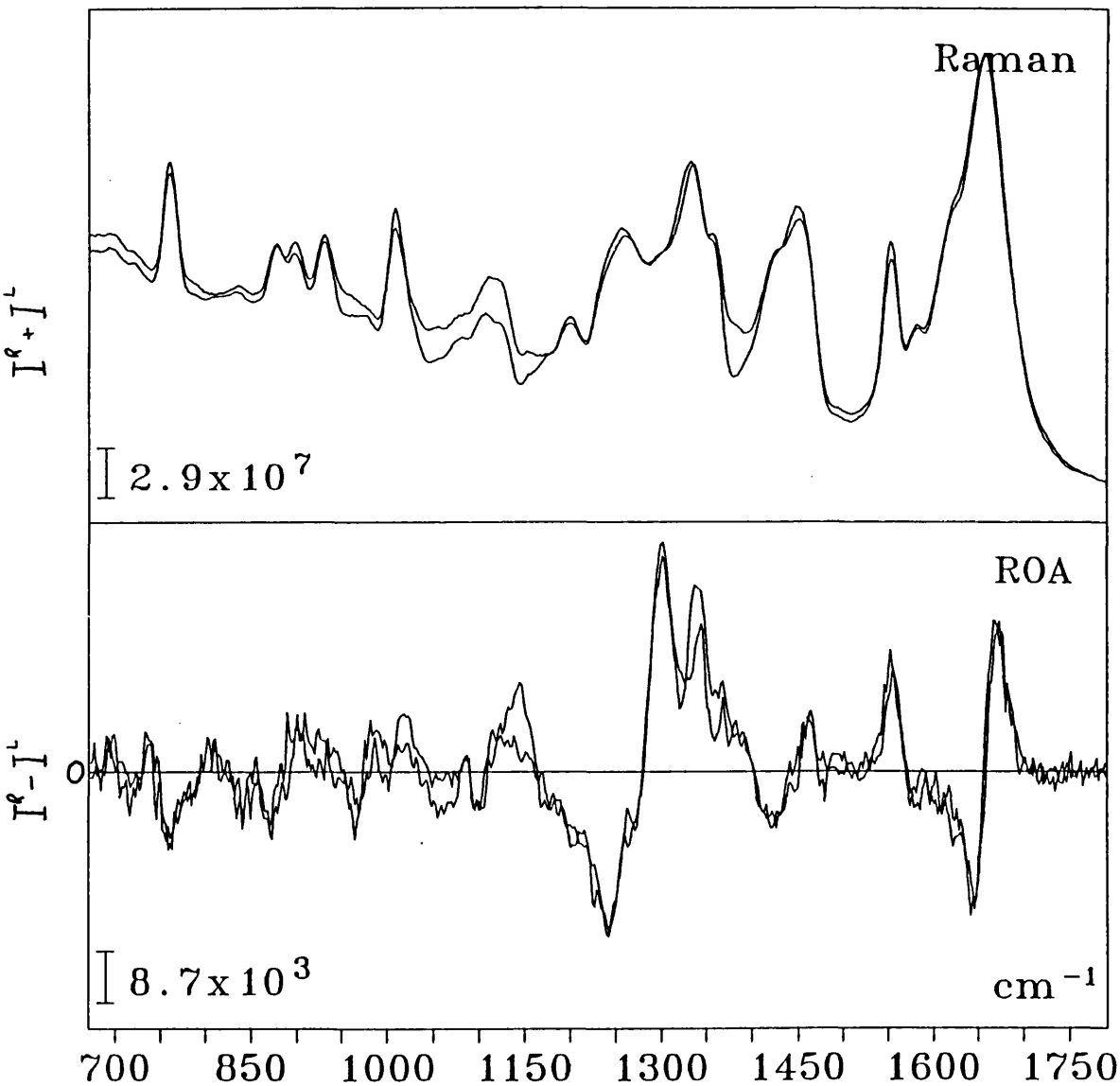


Figure 6.11 The comparison of the Raman and ROA spectra of HEWL-NAG_3 and HEWL.

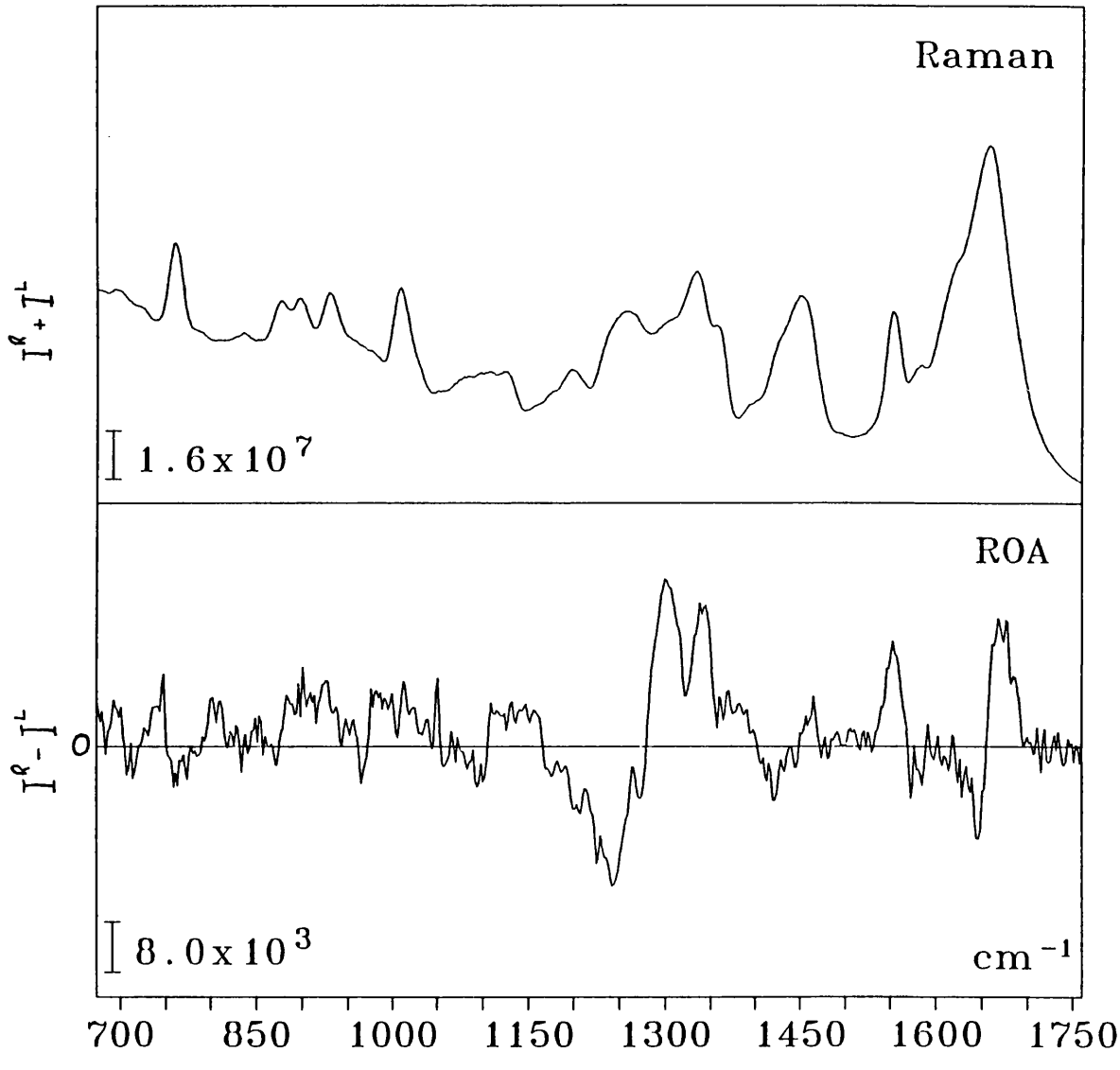


Figure 6.12: The Raman and ROA spectra of HEWL-NAG_3 with the Raman and ROA contributions of NAG_3 subtracted.

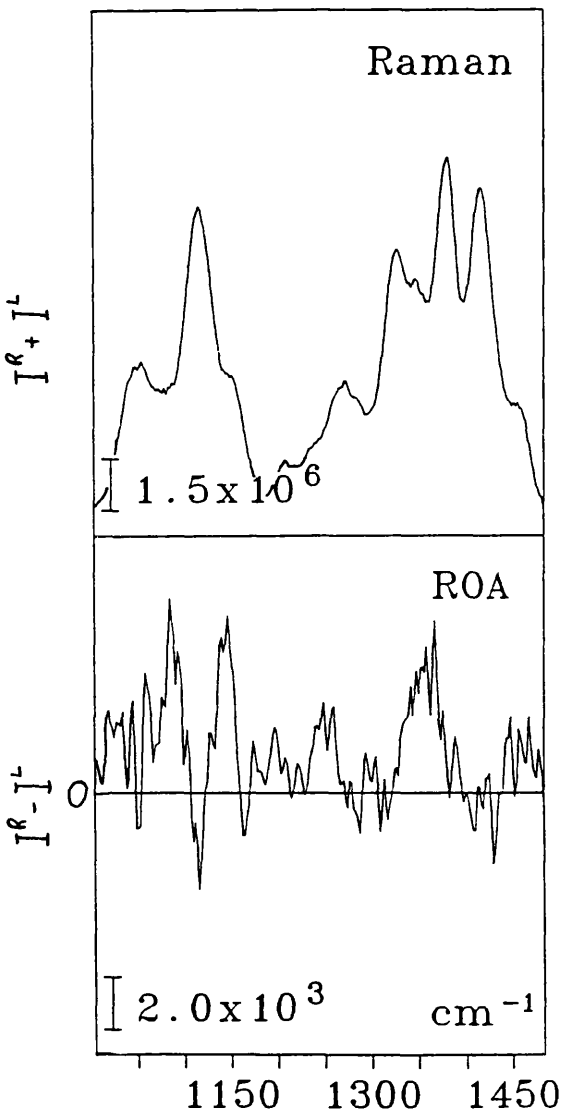


Figure 6.13: The Raman and ROA spectra of NAG_3.

more defined and typically this represents the shielding of tryptophans from the solvent (Yu, 1974; Harada et al., 1986). This effect is also seen in tryptophan hydrogen exchange experiments (Cassels et al., 1978). The 1340 cm^{-1} tryptophan and C-H deformation band moves lower in frequency by a few wavenumbers and the 1255 cm^{-1} amide III Raman band drops in intensity and moves up to 1260 cm^{-1}. A similar change in the amide III Raman has been observed with NAG (Bertoluzza et al., 1992) and ascribed to a 15% increase in α-helix content and a 15% decrease in random coil structure. Although we would not support the final conclusion of these authors it does suggest an interesting possibility as to the cause of the amide III band changes. Inhibitor binding changes to the amide III Raman are the opposite from those that occur in denaturation and these changes may be related to solvent exposure. Tu (1986) proposed that solvent–HEWL hydrogen bonds lower the amide III frequency and this would explain the effects seen here.

Two Raman bands appear in the HEWL-NAG_3 spectrum at 1155 cm^{-1} and 1175 cm^{-1}. These are bands are not seen in the subtraction spectra and so originate from the NAG_3 inhibitor.

The ROA in this region shows some most interesting changes. At 1340 cm^{-1} we see an increase in the positive ROA band upon NAG_3 binding and we see a similar effect in all the ROA spectra of HEWL and HEWL-NAG_3 that we have run. A small part of this band is due to an underlying ROA band from the NAG_3 but, as can be seen from the subtraction spectra, it is not a significant part. There also appears to be changes to the ROA signal at 1330 cm^{-1} and in the 1360 to 1375 cm^{-1} region, but these changes are not seen in some of the earlier ROA experiments. The small, negative 1265 cm^{-1} band moves up to 1270 cm^{-1}. There appear to be changes to the (1245 cm^{-1}) minimum and (1310 cm^{-1}) maximum of the amide III ROA

couplet. In earlier experiments it appeared as if the 1245 cm^{-1} band was sharpening up (becoming more negative), but in the latest set it appears as if the 1310 cm^{-1} band is decreasing in intensity. In both, the changes were quite small. The assignment of these ROA bands is described in chapter 8.

The positive/negative couplet between 1110 and 1160 cm^{-1}, on the low frequency side of the main couplet, appears to become more prominent. The negative component of the couplet becomes more distinct from the main amide III signal and the low frequency, positive component of the couplet becomes far more intense. The subtraction spectra shows that the latter positive feature is prodominantly due to a NAG_3 ROA band.

1150-700 cm^{-1}

The Raman bands at 1125 and 1100 cm^{-1} appear to increase in intensity. However, when we look at the subtraction spectra we see that this increase ios due to NAG_3.

There appears to be a slight change in the ROA between 1000 and 1100 cm^{-1}. However, this change is not seen in some of the earlier ROA spectra. It is not clear from the subtraction spectra (because the signal-to-noise ratio is not high enough) whether this is a real effect or not.

The intensity ratio of the 890 cm^{-1} and 900 cm^{-1} Raman bands alters. These changes are observed in HEWL upon NAG (Bertoluzza et al., 1992) and NAG_3 (Miura et al., 1991) binding. The changes are ascribed to an increase in the low frequency tryptophan band, caused by a change in the hydrogen bonding. The negative ROA band caused by this tryptophan vibration may move to a slightly lower frequency, but the effect is very small.

6.3.5 HEWL hydrogen exchange.

NMR has been used to study the hydrogen exchange kinetics of HEWL. These studies have thrown up a wealth of information and have provided us with a series of simple experiments to follow; we can reproduce the hydrogen exchange experiments and draw parallels between our ROA experiments and the NMR experiments.

Pederson and co-workers divided all the amide links in HEWL into four categories, which depended on their exchange behaviour (Pedersen et al., 1991). The four categories are were defined by a certain level of hydrogen exchange (>90%) within a certain time and under certain conditions. The categories are as follows;

Category I: >90% exchange after 12hrs, pH=4.2, 22°C

Category II: >90% exchange after 25days, pH=4.2, 22°C

Category III: >90% exchange after 25days, pH=7.5

Category IV: <90% exchange after 25days, pH=7.5

As mentioned in section 6.1, the exchange rates of amide hydrogens are dependent upon the environment of the amide; amides that are involved in secondary structure, or are buried, are protected from hydrogen exchange. Although it would be expected that the exchange rate of the amide groups would depend on their dynamic environment, the pattern of exchange remains the same in HEWL crystals, where such dynamic flucuations are unlikely to occur. The secondary structure dependence of the HEWL amides, as quoted by Pedersen and co-workers, is given in Table 6.3.

Table 6.3: The relationship between exchange category and secondary structure.

Exchange category	Secondary structure						Total
	BS	AH	RT	AAHB	ASHB	noHB	
I	0	1	12	4	8	32	57
II	2	8	4	0	3	4	21
III	7	12	7	3	1	2	32
IV	4	10	1	0	1	0	16

The abbreviations are as follows: BS, beta sheet; AH, alpha helix; RT, reverse turn; AAHB, backbone-to-backbone hydrogen bonding; ASHB, backbone-to-sidechain hydrogen bonding; noHB, no backbone hydrogen bonds.

Hydrogen exchange is slow for those amides that are involved in α-helix or β-sheet and those that are buried beneath the protein surface. However, neither factor dominates the exchange rate and a combination of the two is required to rationalize the observed catagories.

These conclusions are those drawn by Pedersen et al., but unfortunately, the ambiguous definition of secondary structure has lead to confusion. In Table 6.3 secondary structure is defined, not by conformational parameters - as they are in McKenzie and White (1991) - but by hydrogen bonding parameters; secondary structure is usually defined by the backbone conformation a particular section adopts (certain Ψ and Φ angles indicate α-helix, or β-sheet etc.); Pedersen et al. have used the hydrogen bonding connectivities of the amide links as their definition (if the amide hydrogen is hydrogen bonded to a carbonyl four residues behind, it is defined as α-helix; if it is three residues behind, it is defined as reverse turn etc.). When we define the sections of secondary structure using Table 6.3 in the introduction the relationship between exchange category and secondary structure becomes different, as outlined in Table 6.4.

Table 6.4: Relationship between exchange category and 'conformationally defined' secondary structure (as defined in McKenzie and White, 1991).

Exchange category	Secondary structure				Total
	α-helix	β-sheet	3_{10} helix	Other	
Type I	14	3	4	36	57
Type II	6	5	2	8	21
Type III	15	5	4	8	32
Type IV	10	5	0	1	16

When we use these definitions of secondary structure the hydrogen exchange-secondary structure relationships become less certain. There appears to be no distinction between the exchange rates of β-sheet, and only a slight distinction exists between the α-helical residues. However, it can seen that most non-secondary structural elements still retain, even within this 'new' definition, rapid exchange.

The positions of the category I and II amides are shown in Figure 6.14.

The hydrogen exchange NMR experiments have defined the rates at which deuteration exchange occurs throughout HEWL. By reproducing this experiment using ROA we can hopefully get a clear idea of ROA-protein conformation relationships.

The hydrogen exchange data also adds to the evidence that the 1340 to 1360 cm^{-1} ROA features originate in the amide III vibrations. This argument was briefly discussed in the assignment of the HEWL ROA, but is given in more detail here. The basis of the argument is that the positive 1340 cm^{-1} ROA signal could come from amide III or tryptophan vibrations. NMR based HEWL hydrogen exchange experiments have quantified the tryptophan exchange rates within HEWL (Wedin et al., 1982). The tryptophan exchange half lives for a 80 mg/ml HEWL solution in acetate

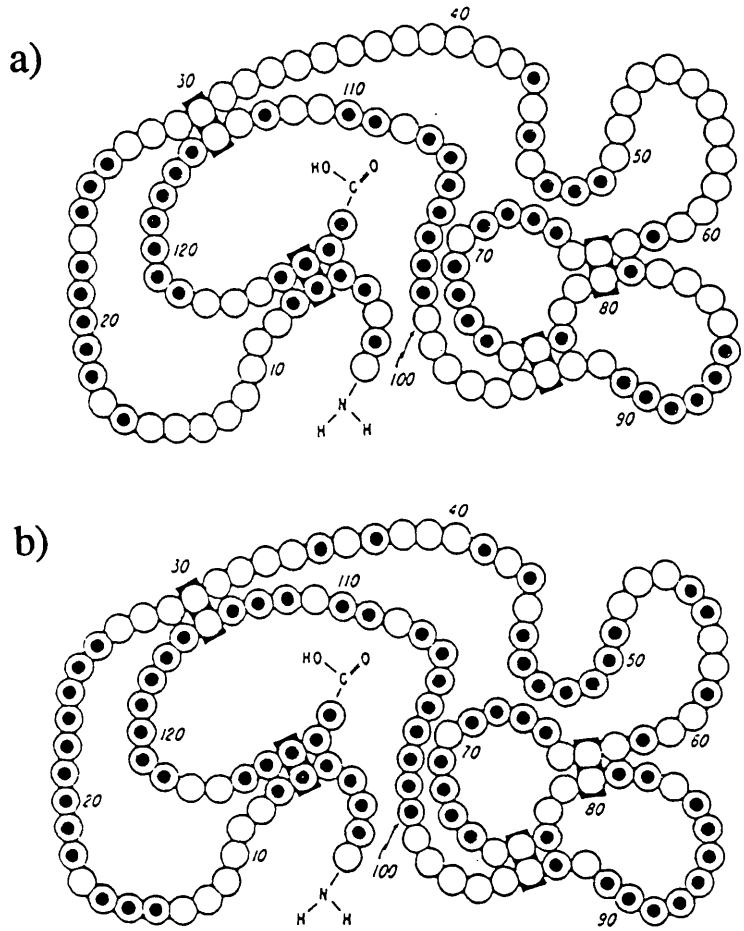


Figure 6.14: The positions of the category I amides (a) and the category II amides (b) within the HEWL sequence are shown by filled circles.

Table 6.5: Tryptophan exchange half-lives.

Tryptophan residue	Exchange half-life
62	Too fast
63	70 sec
108	2 hrs
123	10 hrs
111	10 days
28	3 yrs

buffer (pH=3.8) at a temperature of 21°C are given in Table 6.5. (This table is derived from an Arrhenius plot in Wedin et al., 1982; consequently the values should only be taken as a rough indication of the exchange rate values.) Table 6.5 shows that in the 7 day HEWL sample 70% of the tryptophan residues would have become deuterated; in the 25 day HEWL sample it would be 80%.

No tryptophan ROA spectrum has been obtained in either H_2O or D_2O, but it is known that the 1340/1360 cm^{-1} Fermi resonance tryptophan doublet shows a small (8 cm^{-1}) frequency shift upon deuteration (Harada et al., 1986). If the Fermi resonance was responsible for the positive (hydrated) 1340 cm feature and/or the negative (deuterated) 1340 cm^{-1} feature it would be expected that the largest change in the 1340 cm^{-1} band would occur in the 7 and 25 day HEWL samples. An examination of the partially deuterated HEWL ROA spectra in this region shows that this is not the case.

The preparation of all the deuterated HEWL samples is described in section 6.2.4. The Raman and ROA spectra of the partially deuterated HEWL samples are shown in Figures 6.15 to 6.17.

1750-1500 cm^{-1}

The progression of the amide I band to the (lower frequency) amide I' band can be seen throughout the deuterated series. There is slight a drop in intensity between the hydrated and deuterated samples, probably due to the different solvent rather than the amide I vibrations because there is no obvious change in intensity between the various deuterated samples. The low frequency shoulder on the amide I band, from the aromatic vibrations, gradually disappears as deuteration progresses.

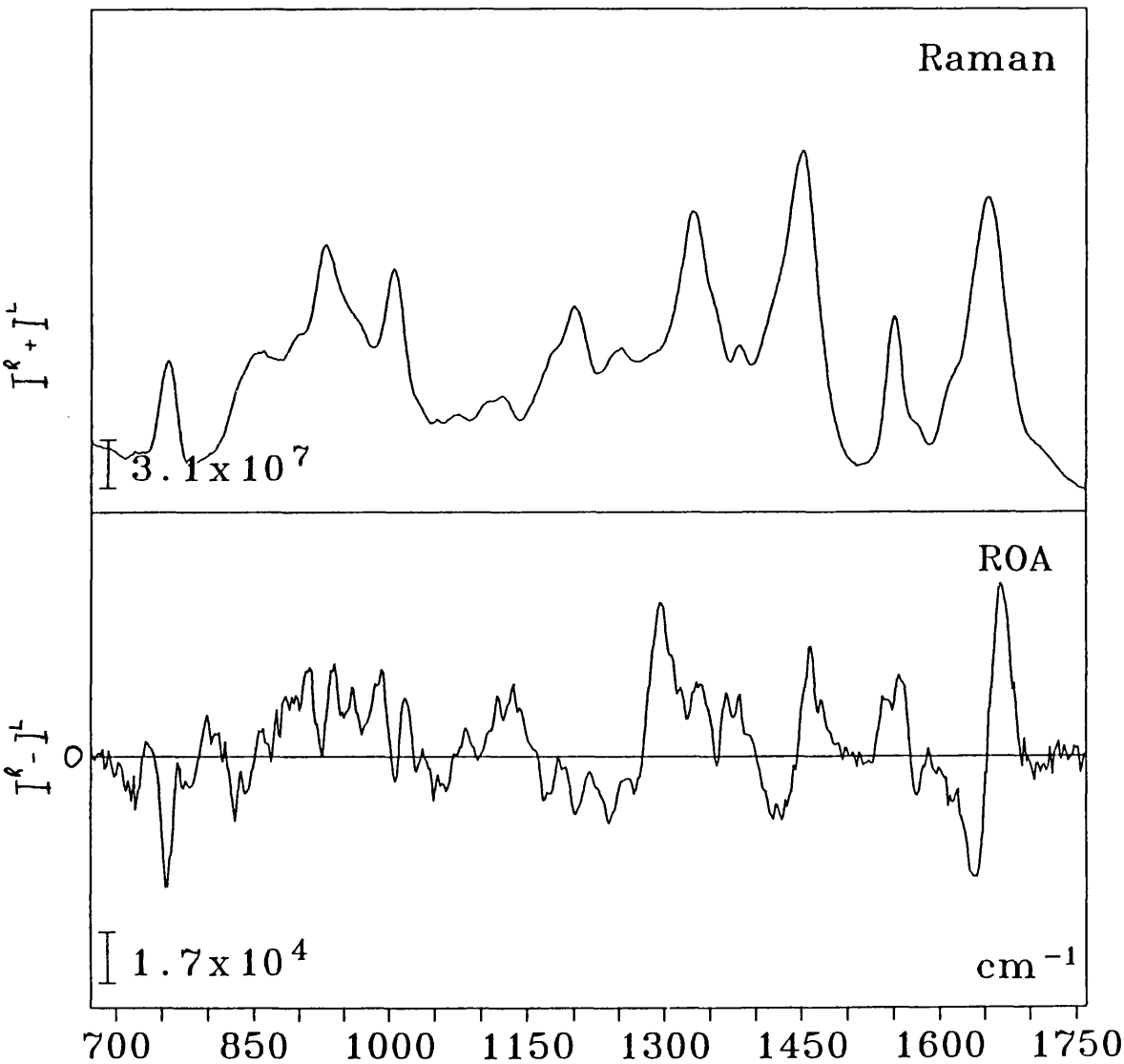


Figure 6.15: The Raman and ROA spectra of the '7 day' HEWL sample.

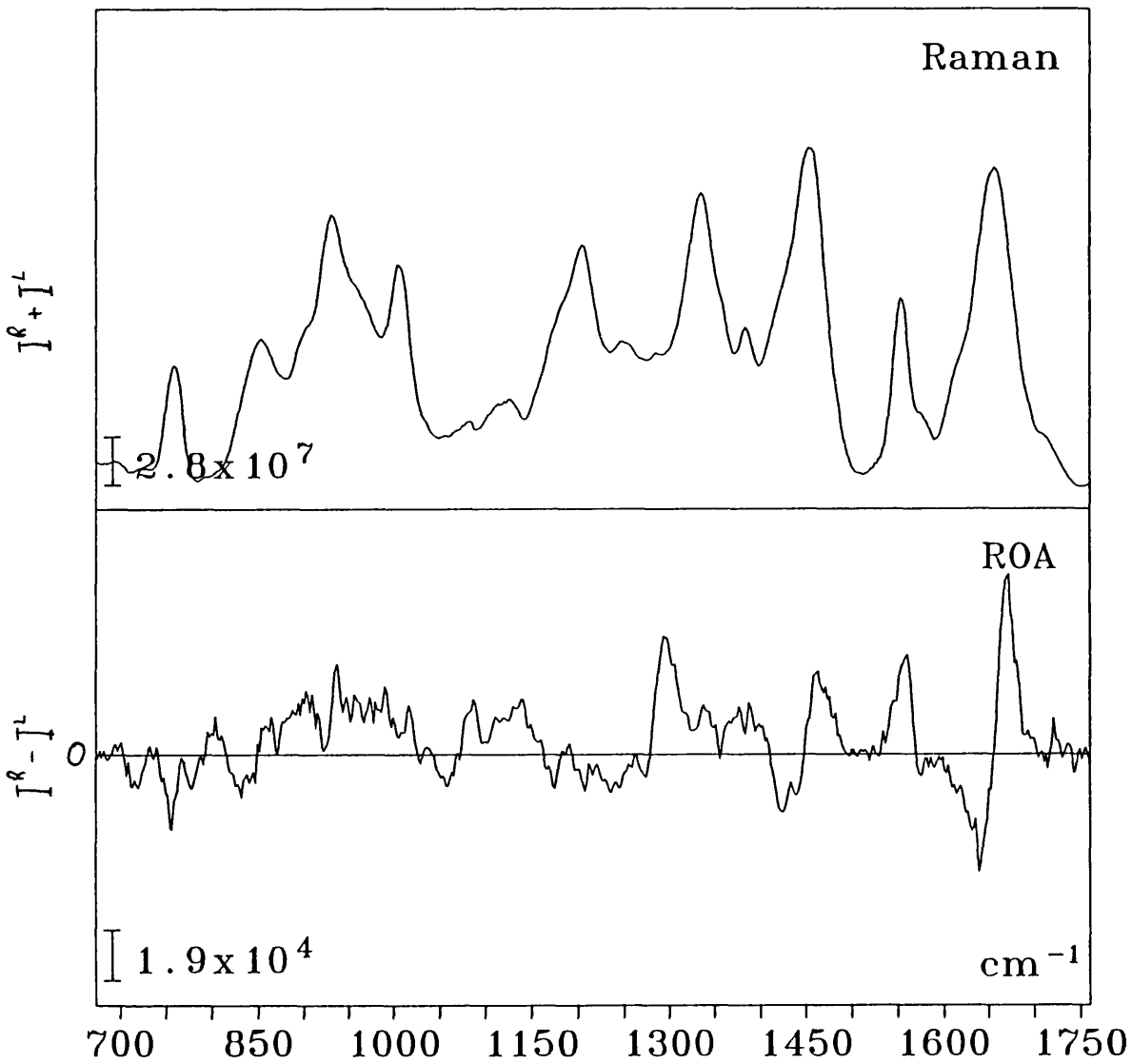


Figure 6.16: The Raman and ROA spectra of the '25 day' HEWL sample.

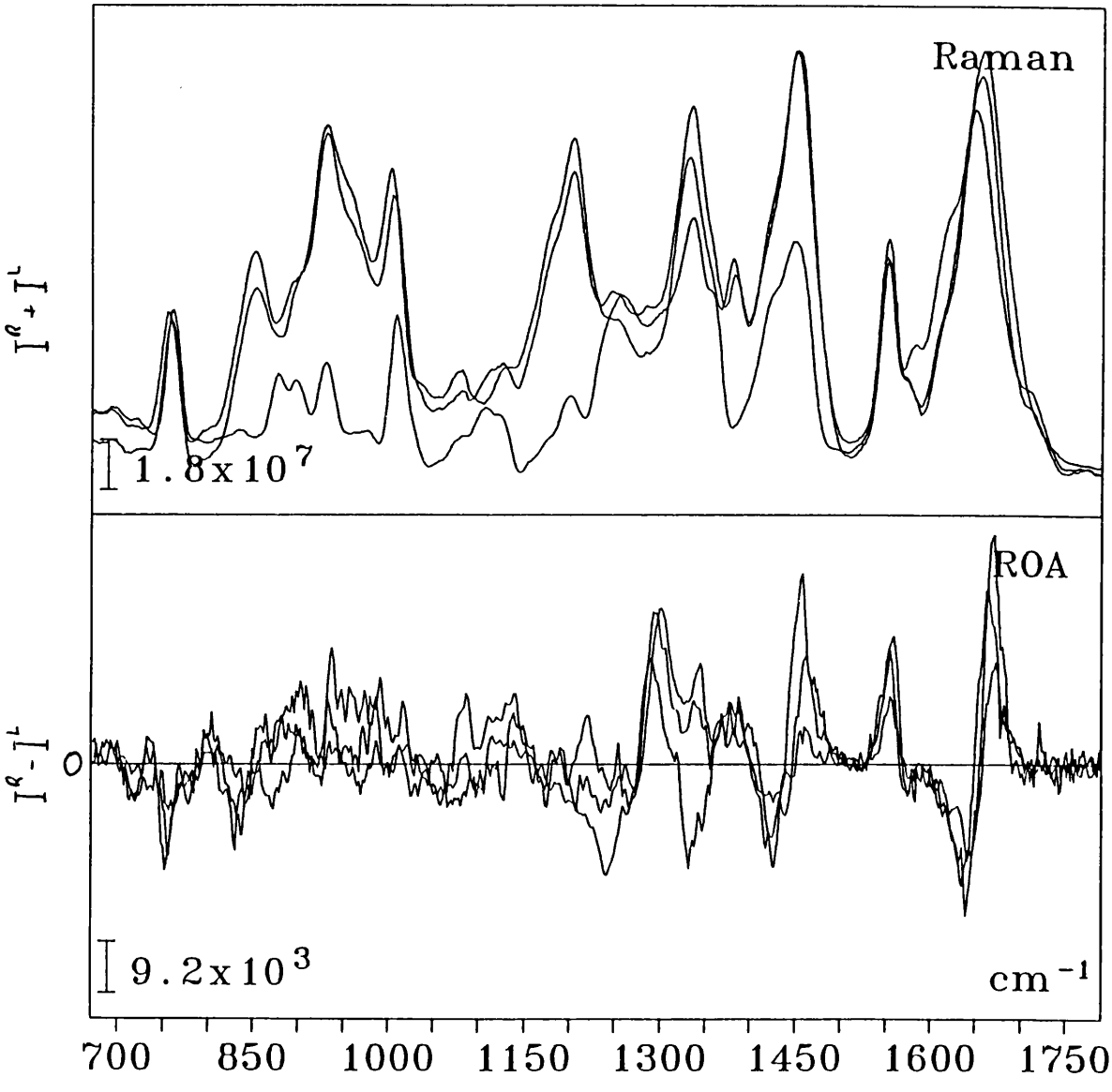


Figure 6.17: The comparison of the Raman and ROA of HEWL, the '25 day' HEWL sample and the fully deuterated HEWL sample.

The amide I/I' ROA of these partially deuterated HEWL samples is more intense than that of the fully hydrated and fully deuterated forms. It appears as if the heterogeneity in the amide groups leads to a more intense signal. The crossover point of the amide I/I' vibrations falls slightly (~3 cm^{-1}) upon partial deuteration. The negative ROA at 1625 cm^{-1} which originates from the aromatic vibrations disappears as deuteration progresses and is obscured by the newly forming amide I' band.

The Raman and ROA signals of the tryptophan vibrations at 1575 cm^{-1} and 1550 cm^{-1} remain the same between the 7 day and the fully deuterated HEWL. The positive ROA 'band' at 1540 cm^{-1} in the 7 day HEWL is an artifact - it is not seen in any of the other HEWL spectra and or in previous deuteration experiments.

1500-1150 cm^{-1}

The next large Raman band is composed of CH_2 deformations, supplemented by the amide II' vibration. The intensity of this band, with respect to the amide I vibration, increases as deuteration proceeds. (Although the '7 day' sample does appear to give an anomalously intense Raman band.) we go from the 25 day HEWL The CH_2/CH_3 deformation ROA remains a couplet as deuteration progresses. As more amide II' vibrations contribute to this band, so the ROA intensity increases.

In the Raman amide III region some of the deuteration trends are quite apparent. The C-H deformation band at 1200 cm^{-1} increases as deuteration progresses - particularly when compared to the C-H deformation band at 1325 cm^{-1}. As the 1200 cm^{-1} band increases so the remaining C-H deformations at 1250 cm^{-1}, which stand on their own in the 7 day HEWL, become a low frequency shoulder in the fully deuterated HEWL.

The ROA of this region demonstrates some very important features about the nature of the amide III ROA. The partially deuterated derivatives allow us to start to plot amide III/structural relationships. Overall, the 7 day HEWL and the 25 day HEWL are quite similar and this is to be expected because only 16% of the amides belong to the type II category alone.

The 1360 to 1400 cm^{-1} broad positive ROA band is seen in the hydrated HEWL ROA spectrum, but as deuteration progresses it gradually becomes more consolidated and intense. The stepwise increase in this band eliminates the possiblity of tryptophan involvment, in the same way as it does for the tryptophan argument. The assignment of this band is discussed above.

The positive 1340 cm^{-1} band remains in the 7 day HEWL and 25 day HEWL and so must originate from type III or IV structural features. In the fully deuterated form there is a strong negative band in this region (high frequency C-H deformations - similar behaviour seen in the peptides). The assignment of the 1340 cm^{-1} band is left to a later point in the discussion. Although it is assumed that the 1340 cm^{-1} band collapses and gives rise to the negative C-H deformations, it is not impossible that the positive band is simply obsured by strong negative C-H deformations which originate from the (hydrated) 1300 cm^{-1} ROA feature.

The hydrated positive and negative ROA bands at 1300 and 1270 cm^{-1}, respectively, remain intact upon deuteration, but do get closer together. The 1300 cm^{-1} band gradually moves to 1290 cm^{-1} and the 1270 cm^{-1} band gradually moves to 1275 cm^{-1}, and their intensity also decreases. These gradual movements indicate that these bands come from C-H deformations from all parts of the structure.

The strong negative 1245 cm^{-1} band vanishes very early on in the deuteration process. The band almost halves in intensity in the 7 day HEWL, and only decreases a little more in the 25 day HEWL and fully deuterated samples. This indicates that the negative 1245 cm^{-1} band is an amide III vibration that originates in exchangable structure within HEWL, probably surface loops and end chains. Similar behaviour is observed, and a similar assignment given, to the 1185 cm^{-1} region of the amide III couplet.

The negative feature at 1210 cm^{-1} (which is part of an 1160 to 1210 cm^{-1} couplet) becomes a positive feature in a gradual fashion. Therefore, this part of the couplet is assigned to general amide III vibrations.

1150-700 cm^{-1}

The Raman band at 1110 cm^{-1} can be seen decreasing in intensity and moving to higher frequency in the partially deuterated HEWL samples. However, the ROA signals in this region are not so clear. A comparison of the fully hydrated, 25 day and fully deuterated HEWL samples indicates that the positive ROA (that is the low frequency component of the 1160 to 1210 cm^{-1} couplet) region moves up in frequency and becomes a sharp band at 1140 cm. However, although not incompatible with the idea, the 7 day HEWL is too noisy in this region to justifiably slot it into this sequence. The changes in the hydrated HEWL 1100 cm^{-1} negative/positive couplet collapses as deuteration progresses and appears to give rise to a positive ROA band at 1075 cm^{-1}. Unfortunately, when deuteration is completed there is no sign of this positive band, but a negative one instead. The latter may have it's origins in the broad, negative 1060 cm^{-1} band of hydrated HEWL. (Similar changes occur in this region in some of the earlier experiments.) The tryptophan Raman band at 780 cm^{-1} gradually move down to 775 cm^{-1} as deuteration progresses (Muria et al., 1991).

The ROA of the rest of this region is difficult to decipher: N-D in plane vibrations underlie this whole region and even in simple peptides the relationship between the hydrated and deuterated features were difficult to determine. However, there are still some Raman features to be discussed. In the 25 day HEWL sample there are high (960 cm^{-1}) and low frequency (900 cm^{-1}) shoulders on the 930 cm^{-1} Raman band. These shoulders are more prominent in both the 7 day and fully deuterated HEWL. This is probably because in the 7 day sample the 930 cm^{-1} band is weaker, which mean that the shoulders are still prominent. In the fully deuterated sample the shoulders become more intense.

Chapter 7

Bovine α-lactalbumin

7.1 INTRODUCTION

7.1.1 Background

A natural extension to the ROA study of HEWL is the study of α-lactalbumin. Alpha-lactalbumin has many of the advantages that HEWL had; it is cheap, commercially available at a suitable purity, is soluble enough and has been well studied by a variety of other techniques. In addition to this, α-lactalbumin has three properties which make a study of it almost mandatory for any new technique: it has a crystal structure almost identical to that of HEWL, it is a metalloprotein and binds very strongly to calcium, and it undergoes a large structural change when the pH drops below 4. Despite α-lactalbumin being such a popular protein for study it remains poorly understood, with all these three interesting features adding to the controversy surrounding the protein and it's properties.

Alpha-lactalbumin, as it name suggests, is extracted from milk (so as a natural consequence is only found in mammals!). Bovine α-lactalbumin (BAL), which was the form of α-lactalbumin used for the ROA study, is 123 amino acids long and has a molecular weight of 14186 amu. The sequence of BAL is shown in Figure 7.1, alongside that of HEWL. Alpha-lactalbumins and lysozymes have about a 40% sequence homology (Acharya et al., 1989). The function of α-lactalbumin is the production of milk during lactation and it combines with an enzyme, galactosyltransferase (GT), to form the lactose synthase complex.

Alpha-lactalbumin alters the specificity of GT from NAG groups on glycoproteins to glucose and in this way promotes lactose synthesis. This is shown diagrammatically in Figure 7.2.

```
                                                10
Glu-Glu-Leu-Thr-Lys-Cys-Glu-Val-Phe-Arg-Glu-Leu-Lys-
Lys-Val-Phe-Gly-Arg-Cys-Glu-Leu-Ala-Ala-Ala-Met-Lys-Arg-His-
                                                10
                              20
Asp-Leu-Lys-Gly-Tyr-Gly-Gly-Val-Ser-Leu-Pro-Glu-Trp-Val-Cys-
Gly-Leu-Asp-Asn-Tyr-Arg-Gly-Try-Ser-Leu-Gly-Asn-Trp-Val-Cys-
                    20                                      30
    30                                              40
Thr-Thr-Phe-His-Thr-Ser-Gly-Tyr-Asp-Thr-Glu-Ala-Ile-Val-Gln-
Ala-Ala-Lys-Phe-Glu-Ser-Asn-Phe-Asn-Thr-Gln-Ala-Thr-Asn-Arg-
                                                40
                                50
Asn-    Asn-Asp-Ser-Thr-Glu-Tyr-Gly-Leu-Phe-Gln-Ile-Asn-Asn-
Asn-Thr-Asp-Gly-Ser-Thr-Asp-Tyr-Gly-Ile-Leu-Gln-Ile-Asn-Ser-
                    50                                      60
        60                                              70
Lys-Ile-Trp-Cys-Lys-Asn-Asp-Gln-Asp-Pro-His-Ser-Ser-Asn-Ile-
Arg-Trp-Trp-Cys-Asn-Asp-Gly-Arg-Thr-Pro-Gly-Ser-Arg-Asn-Leu-
                                                70
                                80
Cys-Asn-Ile-Ser-Cys-Asp-Lys-Phe-Leu-Asp-Asp-Asp-Leu-Thr-Asp-
Cys-Asn-Ile-Pro-Cys-Ser-Ala-Leu-Leu-Ser-Ser-Asp-Ile-Thr-Ala-
                    80                                      90
        90                                                  100
Asp-Ile-Met-Cys-Val-Lys-Lys-Ile-Leu-    Asp-Lys-Val-Gly-Ile-
Ser-Val-Asn-Cys-Ala-Lys-Lys-Ile-Val-Ser-Asp-Gly-Asn-Gly-Met-
                                                100
                                    110
Asn-Tyr-Trp-Leu-Ala-His-Lys-Ala-Leu-Cys-Ser-Glu-Lys-Leu-Asp-
Asn-Ala-Trp-Val-Ala-Trp-Arg-Asn-Arg-Cys-Lys-Gly-Thr-Asp-Val-
                    110                                     120
                              120
    Gln-Trp-Leu-            Cys-Glu-Lys-Leu
Gln-Ala-Trp-Ile-Arg-Gly-Cys-    Arg-Leu
```

Figure 7.1: The sequences of BAL (bold) and HEWL.

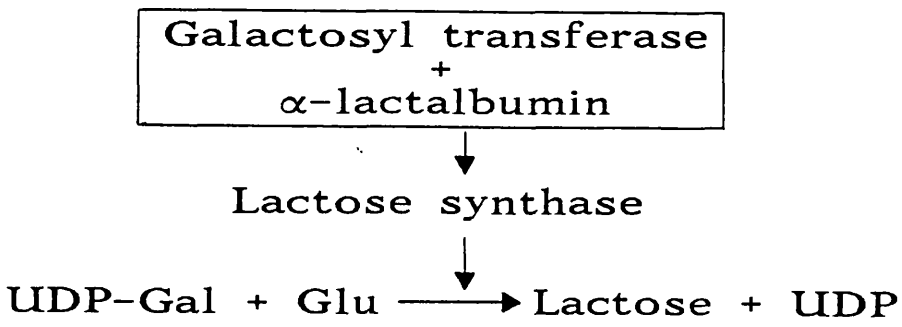


Figure 7.2: The role of GT and α-lactalbumin in lactose synthesis.

The binding between GT and α-lactalbumin is poorly understood and the mechanism by which α-lactalbumin alters the specificity of GT is not known.

7.1.2 Crystal structure of α-lactalbumin.

Although the functions of HEWL and α-lactalbumin are in no way related - they would hardly be classed as homologous proteins - the high sequence similarity lead researchers in the 1960's and 70's to propose models for α-lactalbumin based on the HEWL structure. Although the HEWL structure had been solved in 1965, problems with the crystallization of α-lactalbumin meant that a high resolution structure was not solved until almost 25 years later.

The crystal structure of baboon α-lactalbumin (BabAL) was solved to 1.7 Å resolution by Acharya et al. (1989). Interestingly, this high resolution structure was determined by using HEWL as a starting model (this was justified by previous low resolution structures showing that BabAL was similar to HEWL).

BabAL is 23 x 26 x 40 Å in size and has a structure very similar to that of HEWL and is shown in Figure 7.3. BabAL has an α-helix lobe and β-sheet lobe separated by a cleft. The secondary structural features of α-lactalbumin, as defined by Acharya et al., are given in Table 7.1.

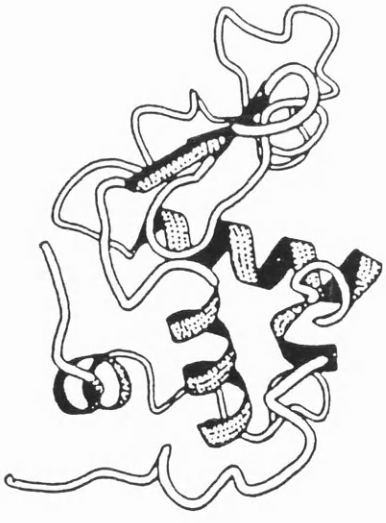


Figure 7.3: The crystal structure of BabAL. (Adapted from McKenzie and White, 1991).

Table 7.1: The secondary structural features of α-lactalbumin.

Structure	Residues
α helix A	5-11
B	23-34
C	86-99
D	105-109
3_{10} helix	12-16
	17-21
	76-82
	101-104
	115-119
β-sheet	40-43, 47-50

The largest deviations between BabAL and HEWL occur in the loop regions of the protein where insertions and deletions in the polypeptide sequence are made. Additionally, significant differences are seen in the C-terminal region.

The crystal structure determined the nature and location of the calcium binding site. The authors observed that the structure of the site was different from the 'EF hand'-type of calcium binding site found in other proteins. The calcium sits in a β-turn like loop between a section of 3_{10} helix and α-helix. (Bearing this structure in mind, and presumably trying to maintain the trend for upper limb analogies, the authors named their new site the calcium binding 'elbow') The calcium binding loop is one of the most ordered parts of the protein (as is demonstrated by low temperature factors in this region). The mainchain conformation of the calcium binding site is very similar to that of the equivalent section of HEWL, but the sidechain conformations are very different. The calcium itself is surrounded by seven ligands, two of which are water. The remaining five protein based ligands are given in Table 7.2.

Table 7.2: The calcium binding ligands from α-lactalbumin.

Residue	Group
Lys 79	Carbonyl
Asp 82	Carboxylate
Asp 84	Carbonyl
Asp 87	Carboxylate
Asp 88	Carboxylate

As with HEWL, α-lactalbumin also possesses a hydrophobic box which is composed of residues Ile 95, Tyr 103, Trp 104 and Trp 60. The cleft is blocked by the Tyr 103, which lies over the α-lactalbumin equivalent of the A and B HEWL binding sites.

The temperature factors (B values) for BabAL and HEWL are compared in Figure 7.4. HEWL is more rigid than BabAL, with the latter showing very pronounced flexibility in the C-terminal region and the first two 3_{10} helices which are both on the surface of the molecule and exposed to the solvent. One pronounced HEWL feature is the increased mobility of the 69 to 74 region, particular when it is compared to the BabAL equivalent which is quite rigid. This region of HEWL undergoes a conformational change when NAG_3 binds (Kumagai et al., 1993; Blake et al., 1981).

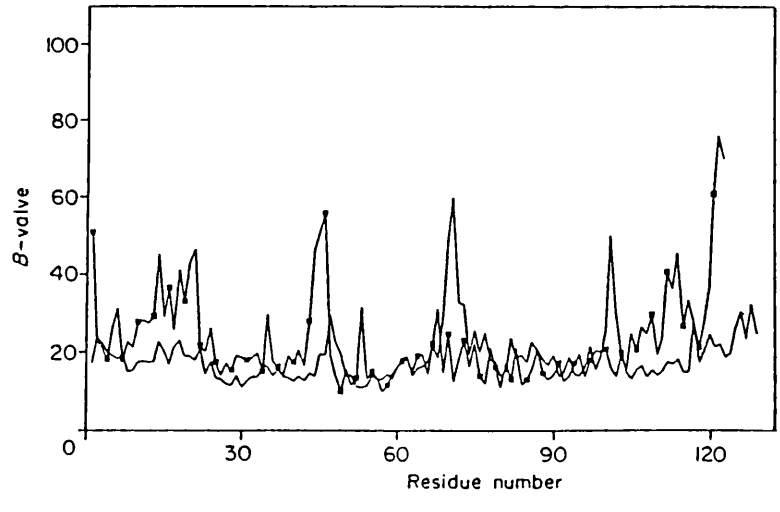


Figure 7.4: The temperature factors of BabAL (■) and HEWL (-). (Adapted from Acharya et al., 1989)

7.1.3 Solution structure of α-lactalbumin.

Although the solution structure of HEWL can be deduced from NMR experiments, this is not yet possible for α-lactalbumin because it is much more difficult to assign the proton resonances of the latter. This difficulty is caused by the poor quality of the NMR signals from α-lactalbumin which are broader then those from HEWL. Alpha-lactalbumin aggregates at the pH's used for NMR experiments and undergoes dynamic fluctuations which occur while the NMR spectrum is being collected, an observation which may be related to the larger temperature factors in the crystal structure of α-lactalbumin. In addition to these experimental difficulties, the assignment of α-lactalbumin is more difficult because of the amino acid sequence; many of the short characteristic amino acid sequences that were used as starting points for the HEWL assignment are not found in α-lactalbumin. NMR studies have found the B and C helices in the native state of guinea pig α-lactalbumin (Alexandrescu et al, 1992). 15 of the 19 identifiable slow exchanging amides in native guinea pig α-lactalbumin belong to helices B and C (Chyan et al., 1993). Much more NMR work has been done on the molten globule state where only the assignable resonances are concentrated on.

The aromatic residues of α-lactalbumin can be assigned, and so NOE experiments can be carried out to measure the distances between the aromatic residues. The distances between the residues as observed by NMR and as deduced from the crystal structure were found to be in good agreement. The solvent exposure of the aromatic residues was determined using photochemically induced dynamic nuclear polarization probes and found to be similar in various form of α-lactalbumin. The conclusion was that the solution structure of α-lactalbumin was similar to the crystal structure (Alexandrescu et al., 1992).

Another technique applicable to the study of solutions and crystals is Raman spectroscopy. The amide III vibrations were traditionally an indication of conformation and using this criteria Yu (1974) proposed that the solution and crystal structures of BAL were very similar. The amide III regions of HEWL and BAL crystals are also very similar.

7.1.4 The low pH, or molten globule state, of α-lactalbumin.

When the pH is lowered to below 4 α-lactalbumin undergoes a conformational transition to a stable, molten globule state. This molten globule state has generated much interest because it has been promoted as a stable protein folding intermediate, or at least an intermediate-like structure which can be easily studied.

The NMR spectrum of molten globule α-lactalbumin shows significant line broadening which represents the interconversion of various different conformations. It has a significant number of chemical shifts which differ from those found in the denatured state, these differing chemical shifts are among those that originate from the hydrophobic core in the native state (Baum et al., 1989).

NMR hydrogen exchange experiments have shown that fewer than 20 amides are slow exchanging in the molten globule state and that many of these amides are also protected in the native state. Some of the protected molten globule amides can be assigned to the B and C helices (Chyan et al., 1993). Inter-residue NOE effects are observed in the molten globule state and this implies that some hydrophobic interations do occur. However, the analysis of chemical shift values and NOE spectra show that the long range tertiary structure found in the molten globule state is non-specific (Alexandrescu et al., 1993). Hydrophobic interactions also appear to play a part in stabilizing the residual secondary structural

elements. The residual hydrophobic interactions and secondary structure elements seem only to be found in the helical lobe of the molecule (Chyan et al., 1993).

UVCD spectra show that the tertiary structure vanishes in the molten globule state. Furthermore, secondary structure estimations from far UVCD spectra show that the secondary structural content varies very little between the native and molten globule states (Chyan et al,, 1993). The difference between the UVCD data and the hydrogen exchange data is explained by transient structures but another possibility is that the UVCD may be detecting localized conformations; that the Φ and Ψ angles of the peptide chain occupy regions of the Ramachandran map that represent α-helix or β-sheet but without the regularity to make up proper secondary structural elements (White, 1982).

The addition of trifluoroethanol (TFE) to the molten globule state of α-lactalbumin generates a different state - the TFE state - which has recently been characterised by NMR, UVCD and fluorescence labelling experiments (Alexandrescu et al., 1994). The pattern of hydrogen exchange protection in the TFE state is similar, but not identical, to that in the molten globule state. The B and C helices are both protected in part and the latter shows an α-helix NOE pattern. An α-helical NOE pattern was found for a section of α-lactalbumin that is not helical in the native state. (The UVCD spectra showed α-helical bands three times more intense in the TFE state than in the molten globule state.) It was proposed that the TFE state was a partially folded state which contained both native and non-native secondary structure, but lacked tertiary structure and was induced by the disruption of the hydrophobic core of the protein.

7.1.5 Metal free (apo) α-lactalbumin.

Alpha-lactalbumin is a metalloprotein which has a very high affinity for calcium, but will also bind other cations. The primary calcium binding site was determined by crystallography and has already been discussed. There is some evidence for the existence of other metal binding sites (possibly up to three) but they could not be observed in the crystal. The role of Ca^{2+} in function of α-lactalbumin is not known. One possible function is the folding of the protein; it has been found that the formation of the correct disulphide bond from reduced α-lactalbumin is dependent on the presence of Ca^{2+} (Rao and Brew, 1989). Calcium also stabilizes the protein against thermal and chemical denaturation. Until recently it was thought that metal binding properties of this type were exclusive to α-lactalbumin, but the discovery that equine and pigeon lysozyme also bind calcium (because of an abundance of aspartate residues in the lysozyme calcium binding elbow) raised new speculations as to the evolutionary and functional links between these two sets of proteins (Aramini et al., 1992; Haezebrouck and Van Dael, 1993).

The metal binding properties of α-lactalbumin have been studied quite extensively. α-lactalbumin will bind a wide range of cations including group I and II metals, transition metals and lanthanides. The estimation of the binding constants appears to be a contentious issue, with different researchers producing widely different results (values for the Ca^{2+}-BAL association constant, K_a, range from 10^6 to 10^{12}).

The issue of cation binding has been a matter of controversy; it has even been proposed that the molten globule state is induced by the dissociation of Ca^{2+} from α-lactalbumin. Early fluorescence data on the binding of Ca^{2+} to BAL showed that removing the Ca^{2+} and lowering the pH produced similar effects on the tryptophan fluorescence spectrum

(Permyakov et al., 1981). Further studies on the various metal binding properties have shown that if Zn^{2+} or Al^{3+} binds to the Ca^{2+} or Mn^{2+} form of α-lactalbumin the latter divalent cations are expelled. Further studies of this aspect of metal binding have lead some researchers to propose the model outlined in Figure 7.5. (A more detailed discussion of the metal binding properties of α-lactalbumin and this suggested model can be found in the recent review by McKenzie and White (1991).)

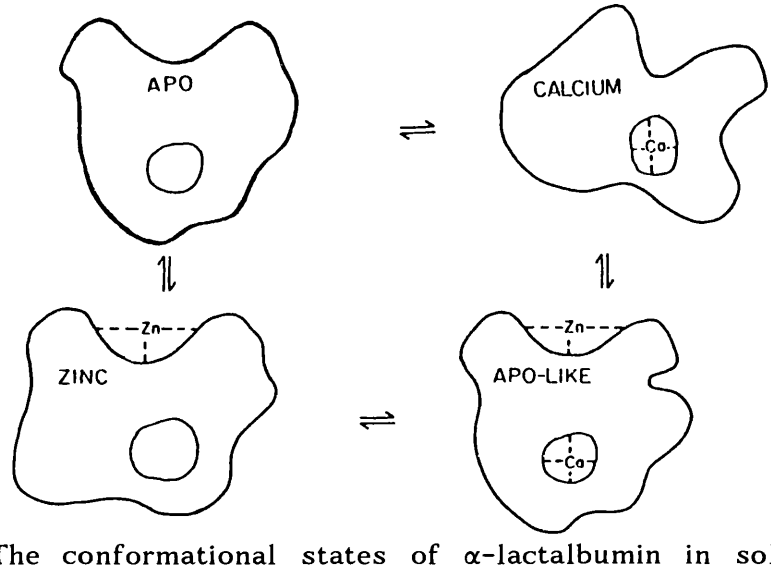


Figure 7.5: The conformational states of α-lactalbumin in solution, and their relation to the binding of various metals.

NMR studies of the metal binding of BAL have shown that the apo structure is similar to the holo structure. Berliner and co-workers reported that several distinct changes in the NMR spectrum of BAL could be observed upon metal binding, but overall the two forms of BAL were "not drastically different in their three dimensional conformations" (Berliner et al., 1987). These authors also noted a change in the surface exposure of some of the aromatic residues. They note that the solvent accessibilty of Tyr 103 and Trp 104 remains the same, but Tyr 18 and Tyr 36 are completely buried upon Ca^{2+} binding. A second NMR study reached the same conclusions and noted that the UVCD as well as NMR showed that the Ca^{2+} free and bound forms were essentially the same (Kuwajima

et al., 1986). This study also proposed that at least two tryptophans (60 and 104) had different environments in the Ca^{2+} free and bound states (Kuwajima et al., 1986). (Kuwajima and co-workers carried out their experiments in a 100mM solution of NaCl, so the results cannot be treated as conclusive for the apo- and holo- forms of the protein.)

Vibrational spectroscopy has also been used to establish the nature of the apo- and holo-α-lactalbumin structures. The amide I' vibrations have been studied with FTIR by Prestrelski and co-workers. Their overall conclusions were that the apo and holo- forms were indeed different (Prestrelski et al., 1991a). However, there are two points about this paper which I wish to address, one of which is dealt with here, the other in section 7.1.6.

The first is the proposal that α-lactalbumin undergoes a slow conformational change when it is dissolved in buffer, the changes being complete in 24 hrs. In deducing this the authors make two assumptions. The first is that the amide I only moves by "1-2 cm^{-1}" upon deuteration and so the shifts larger than this "cannot be explained by deuteration effects alone", and the second is that some amide I bands appear to move to a higher frequency eliminating the possiblity that the effect could be due to deuteration. The first statement is not correct; the value is closer to 5-10 cm^{-1} - as peptide and protein studies show, not only in this thesis, but elsewhere (in particular Urbanova et al, 1991 where amide I and amide I' FTIR spectra of BAL show an 11-13 cm^{-1} shift). The second statement is speculative. The two bands concerned are far smaller than the other amide I bands and may well be due to small high frequency amide I vibrations which were not perceived originally. The authors indicate that the "magnitude of this (conformational) change cannot be estimated due to the confounding effects of the observed deuteration shifts":

although such a conformational change is not impossible, this FTIR data can be more easily explained by deuteration effects. (The ideal test would be to carry out the same experiment on completely deuterated BAL.)

Generally, the UVCD spectra of apo-α-lactalbumin have lead to the conclusion that the apo form has reduced tertiary interactions and is similar to the molten globule state (Segawa and Sugai, 1983). However, careful studies of the effects of buffer have shown that Tris buffer can have a stabilizing effect on apo-BAL, shifting the UVCD spectrum towards that of the holo-BAL (Hiraoka and Sugai, 1985).

7.1.6 Comparison of HEWL and α-lactalbumin.

There have been a variety of studies, including chemical and spectroscopic, that examine the similarities and differences between α-lactalbumin and HEWL. Some of the chemical studies have shown results such as a greater susceptibility of α-lactalbumin disulphides to reduction, a greater susceptibility of α-lactalbumin to trypsin digestion (which suggest a larger tryptophan exposure than is found in the crystal structure), a smaller number of slowly exchanging amide hydrogens in α-lactalbumin (35) than HEWL (44), and a very different stability and pH behaviour of the two proteins.

The spectroscopic evidence usually points to the similarities between the two solution structures. We have already seen that Raman and NMR both propose that the α-lactalbumin solution structure is very similar to the crystal structure. The conclusions from UVCD spectra appear to vary. Early researchers claimed that the UVCD curves of BAL and HEWL were indistinguishable, but later work appears to have contradicted this. In particular some UVCD data comparing various proteins found that the UVCD curve of human α-lactalbumin was similar to that of human

lysozyme, but different to that of HEWL. The crystal structures of all these proteins are quite similar (Acharya et al., 1991). The most recent UVCD study was done as part of a VCD discussion by Urbanova and co-workers: their conclusions were that BAL and HEWL had different solution structures.

The disparity between the chemical and spectroscopic conclusions has been explained both by structural (i.e. the solution conformation of the two proteins is indeed different) and dynamic arguments. The dynamic argument is that the α-lactalbumin exists as an equilibrium between a tightly folded conformation and an open, or loose, conformation thus explaining the chemical reactivity of α-lactalbumin. The HEWL, on the other hand, simply exists in one tightly folded conformation.

Two of the most recent vibrational spectroscopy papers have arrived at different conclusions about whether α-lactalbumin and HEWL have a different solution conformation or not. The first is an FTIR paper and the second is a VCD paper. Both of these papers deserve to be commented on because they are very close to the subject of this work.

A comparison of HEWL and BAL was made by examining the FTIR spectra of the amide I' spectra of BAL and HEWL (Prestrelski et al., 1991b). Based on this work they concluded that BAL and HEWL have "a high degree of correspondence in the secondary and tertiary structure". However, differences do exist in the amide I' region and these differences were large enough to provoke Urbanova and co-workers to comment that "this claim (that the proteins had the same conformation) was made despite the fact that the overall band shapes of the two protein amide I' spectra are noticeably different" (Urbanova et al., 1991). By examining the amide I' results of HEWL, holo-BAL and apo-BAL it can be shown that this objection is valid. A comparison of these three spectra show that

although the differences between the bands in apo-BAL and holo-BAL (a combined total of 21 cm^{-1} and 0.32 intensity units) are greater than those between holo-BAL and HEWL (a combined total of 15 cm^{-1} and 0.23 intensity units), it is not entirely clear why the former differences imply a different protein structure, and the latter imply a similar protein structure.

The VCD spectra of HEWL and BAL are very different (Urbanova et al., 1991) and an analysis of the BAL VCD spectrum indicates that BAL is very similar to proteins that have a high β-sheet or "other" structural content, but have very little α-helix. However, the BAL VCD spectrum becomes similar to that of the the HEWL when BAL is solvated in a 33% propanol solution. These results are also backed up by FTIR results and UVCD spectra. This is explained by the α-helical inducing properties of short chain alcohols.

Short chain alcohols have two effects on proteins. First, they act on the conformation of proteins and peptides. They are known to stabilize dynamic peptide molecules (Alexandrescu et al., 1994), particularly helical regions of peptides (Buck et al., 1993) and several β-sheet proteins have been observed to adopt α-helical conformations in alcohol solutions (Buck et al., 1993 and references contained therein).

Secondly, alcohols will denature proteins and this is classically shown by the effect that alcohols have on the thermal melting point of HEWL (Velicelebi and Sturtevant, 1979). In order to check the affect propanol has on BAL a series of DSC experiment were run. The experimental conditions of the VCD experiment were mimicked (exept for protein concentration). The experiments were run by Mr. S. Robertson and an outline of the theoretical and practical aspects of the DSC experiment can be found in Cooper and Johnstone, 1994.

The results are given in Table 7.3, and are graphed together with the parallel results from HEWL in Figure 7.6.

Table 7.3: Thermal melting point (T_m) of BAL against propan-2-ol concentration (w/w).

Propan-2-ol concentation (%)	T_m (°C)
0	59.5
20	29.0
33	not observed

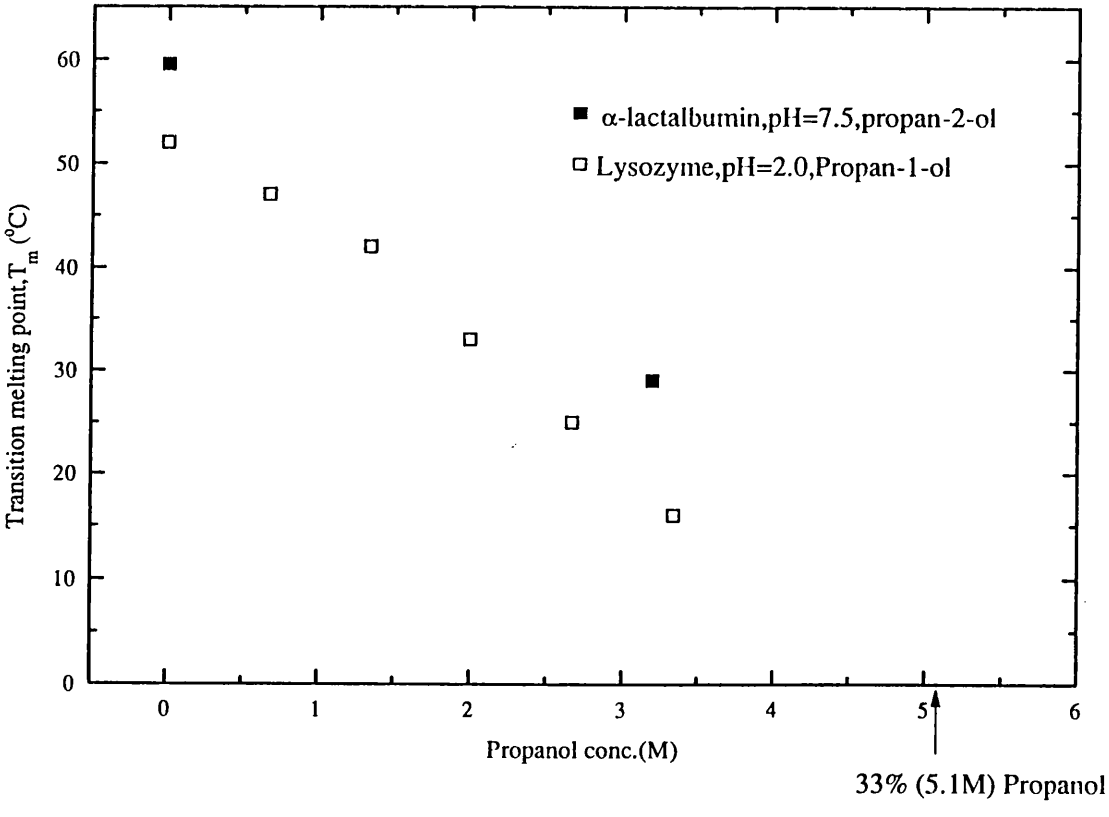


Figure 7.6: T_m of BAL and HEWL versus propanol concentration (M). HEWL results taken from Velicelebi and Sturtevant (1979).

It can be seen from the graph that in 33% propanol the BAL is denatured at room temperature, and this is why a T_m value is not observed in this solvent. Propanol acts as a chemical denaturant and so it

is only at a certain concentration that the BAL will appear denatured. This effect is in fact seen by Urbanova and co-workers where they note that "the presence of up to 20% propanol.. does not change any of the spectral band shapes..but a sudden changes is seen..in a solution with 33% propanol added".

These results show that the BAL in 33% propanol (the 'PrOH-state') is a chemically denatured state, a point not considered by Urbanova et al. Neither the PrOH-state nor the molten globule state have measurable T_ms (Haynie and Freire, 1993). It should be noted that the DSC and VCD experiments monitor different manifestations of protein unfolding; the DSC tells us about the thermodynamics of the 33% propanol BAL, the VCD tells us about the structure. Consequently, the VCD results are not wrong, but as it stands they imply that the BAL structure is only similar to the native HEWL structure in a denatured, or possibly molten globule like, state.

7.2 EXPERIMENTAL METHODS

7.2.1 Bovine α-lactalbumin.

The BAL samples were purchased from Sigma. The native and molten globule BAL studies were done using Grade I BAL which contained a 2 molar excess of Ca^{2+}. The calcium free studies were done using a specially decalcified form of Grade III BAL. According to Sigma literature Grade III BAL contains 0.2 moles of calcium per mole of BAL. The BAL concentrations were measured using UV spectroscopy. The extinction coefficient, E_{280nm}, of BAL was taken as 2.01 $mlmg^{-1}cm^{-1}$ (Gill and Von Hippel, 1989). The extinction coefficient of apo-BAL was taken as the same value, although in reality it is 1.95 ml $mg^{-1}cm^{-1}$ (Yutani et al, 1992).

The BAL samples were prepared the same way as the HEWL samples (see section 5.2.1). The Ca^{2+} bound BAL is soluble over a wide pH range, and is suitably soluble in sodium acetate buffer at pH=4.5 or sodium borate at pH=8.0. For the Ca^{2+} bound and free experiments the higher pH was selected because Ca^{2+} free BAL is not soluble at lower pH's. The molten globule state was prepared by using a 100 mM glycine buffer (pH=1.51).

7.2.2 Calcium free BAL.

The Grade III (Ca^{2+} free) BAL provided by Sigma contains a small amount of Ca^{2+}, and possibly other cations from buffer salts. Therefore it was decided to try and remove the cations before analysis. The experimental procedures for this was as follows:

1) A 20mM EDTA buffer solution was made up and titrated to pH=8.2 using 33% ammonia.

2) Approximately 100mg of Sigma Grade III BAL was dissolved in 10mls of the EDTA/NH_3 buffer. This solution was stirred for 1 hour.

3) This solution was dialysed against 5x200mls volumes of the EDTA/NH_3 buffer, then against 7x3l volumes of deionized water.

4) The dialysed solution was lyophilized.

The resulting apo-BAL was reasonably soluble (75 mg/ml) in sodium borate buffer (pH=8.0) but not so in Tris buffer (45mg/ml at pH=8.0).

7.3 ASSIGNMENTS OF THE RAMAN AND ROA SPECTRA OF BAL.

This section parallels the equivalent HEWL section (6.3). Therefore it is laid out in a similar fashion, but for the sake of brevity I have not repeated information which is stated in the previous section; in the following assignments I have ascribed bands to residues, or sidechains, but the precise origin of the Raman vibration within the structure is not repeated when it has been previously discussed in the HEWL section.

7.3.1 Native α-lactalbumin.

The Raman and ROA spectra of native BAL are shown in Figure 7.7.

1750-1500 cm^{-1}

As with HEWL, there are four bands in this region. The large 1655 cm^{-1} Raman band is the amide I vibration combined with the H_2O bending. This is about 10 cm^{-1} lower, and slightly broader, than the amide I vibration of HEWL, and although these changes are not recorded in previously published Raman spectra (Yu, 1974) they are in accord with the frequency shifts seen in recent FTIR spectra (Urbanova et al., 1991). (The FTIR spectra in H_2O show that the native state BAL has two amide I peaks at 1650 and 1660 cm^{-1}, whereas native HEWL has one main peak at 1657 cm^{-1}.)

The next band at 1620 cm^{-1} belongs to aromatic residues, and the next two bands at 1580 and 1550 cm^{-1} originate from tryptophan residues. The intensity of these tryptophan bands is less in BAL than in HEWL because the former contains only four whereas the latter has six (McKenzie and White, 1991).

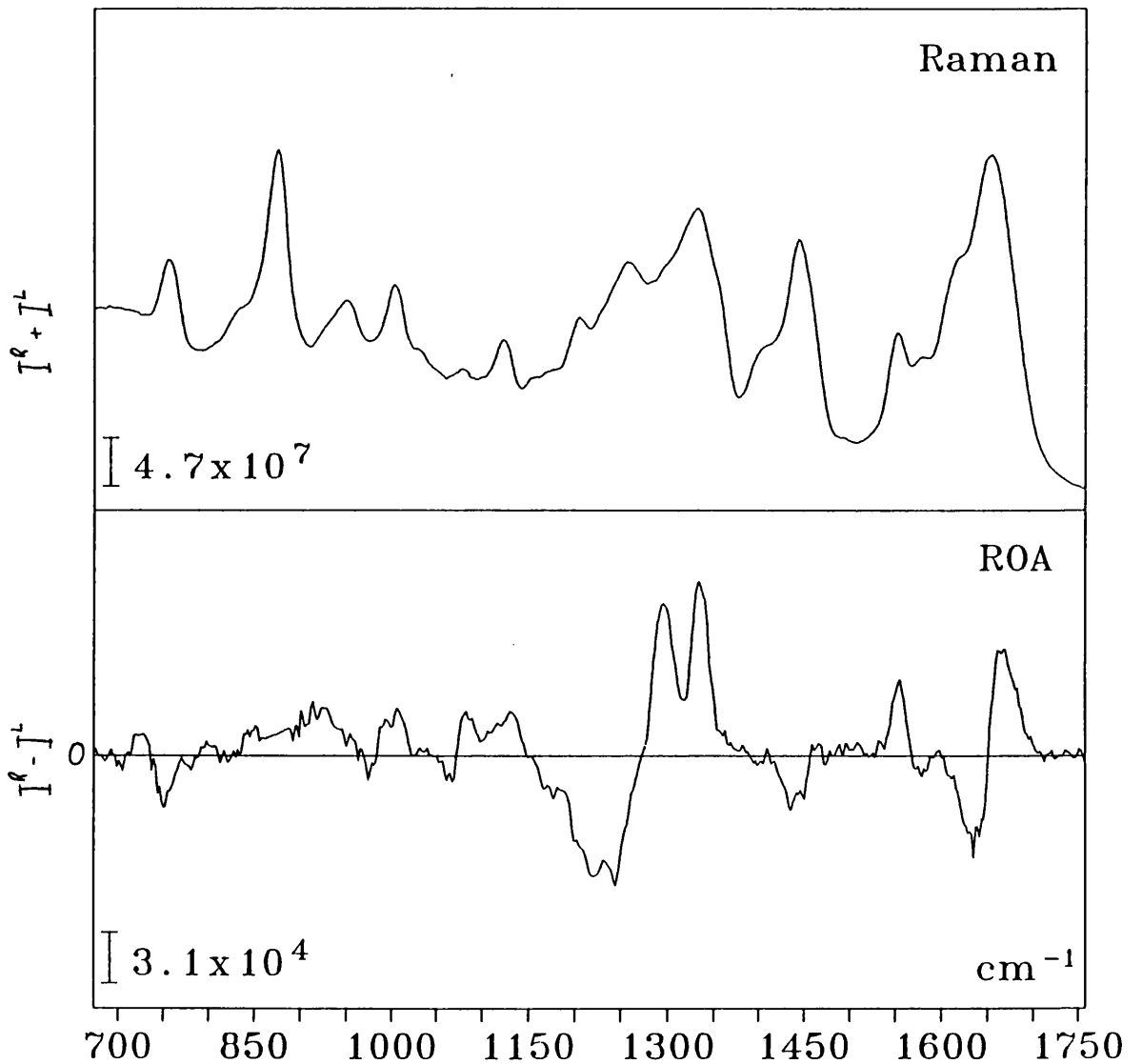


Figure 7.7: The Raman and ROA spectra of native BAL.

The ROA of the amide I vibration is very similar to that found in HEWL; the negative/positive couplet is evenly balanced and is of a similar intensity to that found in HEWL. As with the Raman amide I, the ROA couplet appears to be a little broader in BAL. (This may be responsible for the disappearance of the negative aromatic 1625 cm^{-1} ROA which is seen in HEWL.) The positive/negative tryptophan couplet between 1550 and 1580 cm^{-1} is very similar in BAL and HEWL, but may be slightly better better defined in the former.

1500-1150 cm^{-1}

The narrow Raman CH_2 deformation band appears at 1445 cm^{-1} in BAL, which is slightly lower and sharper than in HEWL. The tryptophan vibrations which form the low frequency shoulder to this band are also lower than they were in HEWL, at about 1410 cm^{-1}. This 20 cm^{-1} drop may be related to the hydrogen bonding of the tryptophan residue which has been shown to alter the frequency of this band (Muira et al., 1989). For the relationship that has been proposed by Muria and co-workers to be upheld it would mean that the BAL tryptophans are weaker hydrogen-bonded than those in HEWL.

The ROA of this CH_2 deformation and tryptophan band is negative, as has be observed in previous BAL ROA spectra (Barron et al., 1992a). It is not clear why the CH_2 deformation Raman and ROA bands are different in BAL than in HEWL. The difference probably represents a change in the sidechain composition between the two proteins rather than a secondary or tertiary structural change. This is indicated because the Raman and ROA CH_2 deformation band does not alter when the molten globule state is generated.

The next large Raman band at 1340 cm^{-1} is a combination of the two tryptophan Fermi resonance bands and high frequency amide III/C-H deformation vibrations. The 1360 cm^{-1} tryptophan band cannot be distinguished from the main band in BAL as it could in HEWL, but is observed at higher resolution (Yu, 1974). In that study the 1360 cm^{-1} band was reported as solvent sensitive. Tryptophan studies back up this relationship, but point out that the C-H vibrations will interfere in proteins (Harada et al., 1986).) Both the 1340 cm^{-1} Raman band and the 1255 cm^{-1} amide III Raman band are of a similar frequency and intensity in BAL and HEWL, although they are both broader in the former. There is no evidence of this in the previous Raman comparison (Yu, 1974). To the low frequency side of the amide III band there are three smaller bands. The 1200 cm^{-1} Raman band originates from phenylalanines and tyrosines. It is slightly more intense in BAL than in HEWL and this reflects the increased occurrence of these residues in BAL. (BAL contains four tyrosines and four phenylalanines, HEWL contains three tyrosines and three phenylalanines.) The 1160 and 1180 cm^{-1} bands are either aromatic residue vibrations, or conformationally insensitive C-N stretches.

The amide III Raman spectra of HEWL and BAL crystals are very similar (Yu, 1974). Our own HEWL and BAL solution Raman spectra also indicate a similar structure. Although the validity of structural predictions from the amide III region has recently been called into question by normal mode calculations (Krimm and Bandekar, 1986), if this argument were to be applied to the HEWL and BAL comparison then it would have to be postulated that the sidechains alter the amide III mode in such a way that it became very similar to what it was before.

The broad amide III ROA couplet has five features superimposed on it. Two strong positive features appear at 1340 and 1295 cm^{-1}. These two

features bear a strong resemblence to the two positive features in HEWL. However there are differences, the most significant being that the 1340 cm^{-1} band is far larger than the equivalent in HEWL. Two negative bands appear at 1240 and 1220 cm^{-1} of which only the former is found in HEWL. A positive/negative couplet, similar to that in HEWL, appears between 1130 and 1190 cm^{-1}. All these features are reproducible and are seen in previous BAL spectra (Wen, 1992; Barron et al., 1992a). The assignment of these bands is left to a special amide III discussion section (chapter 8).

1150-700 cm^{-1}

The first two Raman bands in this region are at 1130 and 1080 cm^{-1}. These bands are very similar to those found in HEWL; they are the same in the molten globule state and deuterated BAL (as those in HEWL were insensitive to deuteration and denaturation). Consequently these bands are more likely to be dominated by sidechain vibrations (methyl rocks for the 1130 cm^{-1} band and aromatic C-H vibrations for the 1080 cm^{-1} band).

The ROA of this region shows that the methyl rock vibrations generate a positive ROA feature, as they did in HEWL. This positive feature remains in the molten globule state. At 1100 cm^{-1} the ROA intensity drops, whereas in HEWL a couplet was found. At 1080 cm^{-1} there is a negative/positive couplet. This probably originates from the aromatic C-H bends, as with the HEWL again, but with BAL the ROA couplet is more prominent and narrower. The main difference between BAL and HEWL is the absence of the negative 1100 cm^{-1} band in the former. This negative band, and it's associated couplet, originated from the HEWL backbone vibrations which also gave a Raman band. There is no sign of the Raman band in BAL and so the absence of the ROA signature is not surprising.

The 1005 cm^{-1} Raman band originates from tryptophan and phenylalanine residues. This band is 10 cm^{-1} lower than in HEWL. The ring stretch frequency for phenylalanine is 1004 cm^{-1} and for tryptophan is 1016 cm^{-1}: a possible explanation for the 10 cm^{-1} drop in frequency between the two proteins is the ratio of phenylalanines to tryptophans in the two proteins (3:6 in HEWL and 4:4 in BAL). As with HEWL the high frequency shoulder on this band, at 1030 cm^{-1}, is from phenylalanine residues. The positive ROA feature at 1000 cm^{-1} is associated with the ring stretch vibrations.

There is an asymmetric band at 950 cm^{-1} which probably involves C-C stretching vibrations. The ROA associated with these vibrations is broad and positive, more so than in HEWL where the positive bands were quite small. The strong band at 870 cm^{-1} originates from the borate buffer, the low frequency shoulder at 830 cm^{-1} is the tyrosine Fermi resonance and it appears to be a lot more intense in BAL than in HEWL. Previous Raman spectra of these proteins in water show that the buffers do not exaggerate this tyrosine band (Barron et al., 1992a). The borate band probably obscures another two Raman bands at 850 and 880 cm^{-1}, which are seen in Raman spectra of BAL in water (Wen, 1992). The indole ring breathing vibration of tryptophan comes at 755 cm^{-1} and it has negative ROA associated with it as in HEWL.

7.3.2 Molten globule BAL.

The Raman and ROA spectra of molten globule BAL are shown in Figure 7.8.

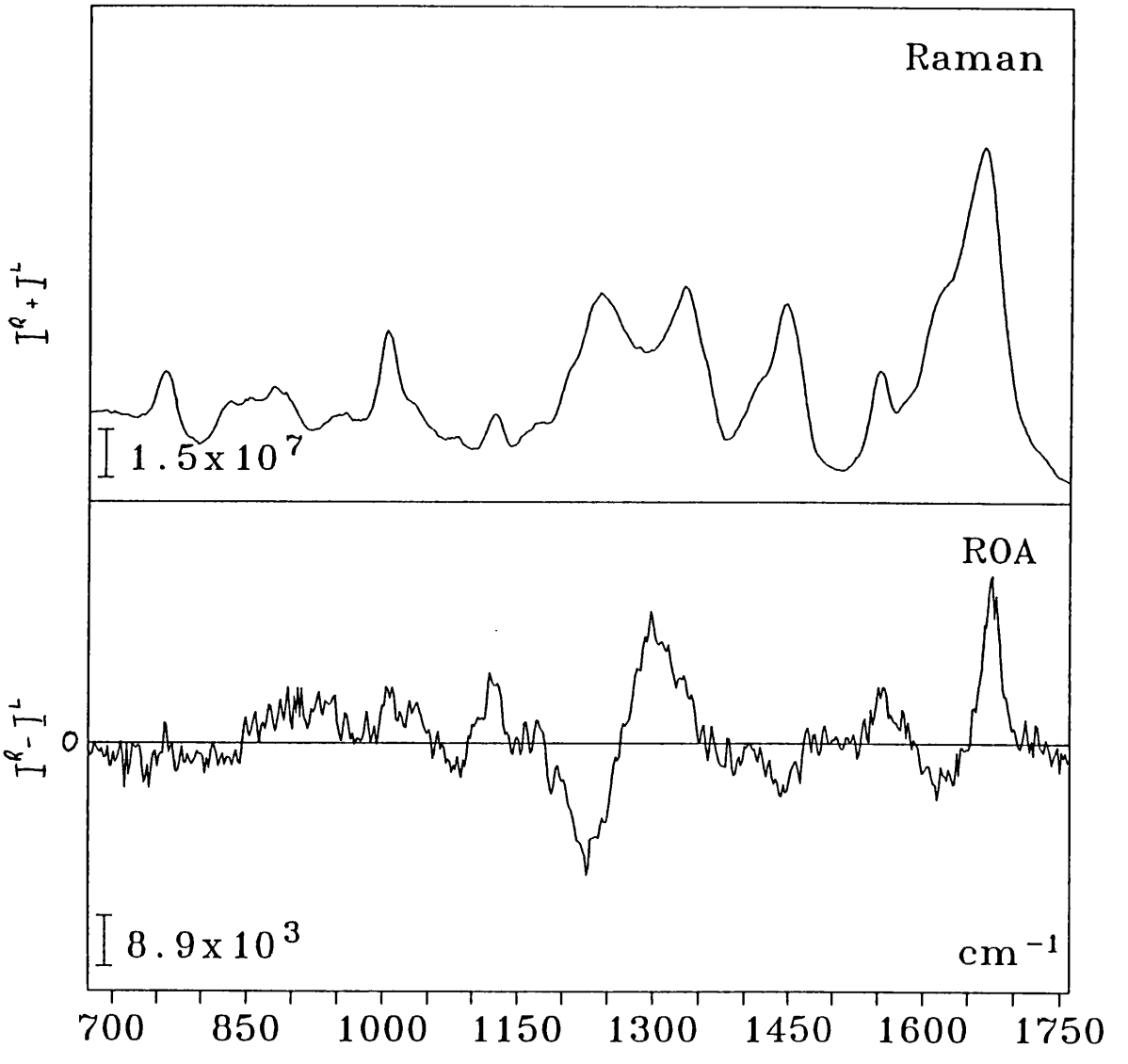


Figure 7.8: The Raman and ROA spectra of molten globule BAL.

1750-1500 cm^{-1}

The amide I of BAL has increased in frequency by 15 cm^{-1}, to 1665 cm^{-1}, upon induction of the molten globule state, whereas HEWL only increased by 5 cm^{-1} upon complete denaturation. The molten globule state of BAL has a slightly narrower amide I Raman band at 1665 cm^{-1} and is very similar in shape and frequency to the amide I of HEWL. Urbanova and co-workers (1991) also found similarities between the amide I' FTIR spectra of the PrOH state of BAL and the native state of HEWL. This raises interesting questions as to the nature of the PrOH state because there appears to be a spectroscopically observed correlation between the PrOH and molten globule states and native HEWL. There may be a parallel between the PrOH state and the molten globule-like TFE state which NMR has characterised (Alexandrescu et al., 1994). The comparison between these different states is limited by the lack of high resolution spectra of the native, molten globule, TFE, PrOH and denatured states of BAL using the same experimental technique. Recent amide I VCD spectra show that the TFE and PrOH state of BAL are very similar, showing a intense characteristic 'W' pattern. Molten globule BAL (pH=2) shows a less intense 'W' pattern (Keiderling et al., 1994).

The amide I ROA of the molten globule state is significantly different from that of the native BAL. Although the amide I is still a couplet, and the crossover frequency of that couplet is the same as it was in native BAL, the negative portion of the couplet decreases significantly. Although, this positive intensity coincides nicely with the recent suggestion that a positively-biased amide I ROA couplet originates from α-helix (Wen et al., 1994a; Barron et al., 1994), recent ROA experiments have shown that this couplet is more balanced and is similar to that found in denatured HEWL.

The negative ROA aromatic vibrations at 1620 cm^{-1} remain as they were in native BAL but become more prominent. This occurs partly because of a drop (~5 cm^{-1}) in the frequency of the ROA (as occurred in HEWL upon denaturation) and partly because of the changes to the amide I ROA. The tryptophan vibration at 1580 cm^{-1} changes both its Raman and ROA signals and this reflects the environmental changes which occur at the tryptophans. The Raman and ROA of the 1550 cm^{-1} tryptophan band remains the same in both the native and molten globule BAL.

1500-1150 cm^{-1}

The Raman and ROA signals of the CH_2 deformation band at 1445 cm^{-1} remains the same in molten globule BAL as they were in native BAL. The tryptophan Raman band goes up by 10 cm^{-1} to ~1420 cm^{-1} in molten globule BAL and this increase reflects again the change in the tryptophan environments when pH is lowered. Studies on this frequency increase indicates that the tryptophans are more strongly hydrogen bonded (Miura et al., 1989).

The 1340 cm^{-1} Raman band is narrower in the molten globule state that it was in the native BAL spectrum. This is because of the change in the C-H deformations or high frequency amide III vibrations which underlie the main tryptophan vibration. Although this is not immediately clear from our Raman spectra Yu's higher resolution data show that the tryptophan band is still clearly present at low pH. Yu also reports that the 1360 cm^{-1} tryptophan band is not seen in the molten globule state and tentatively ascribes this to the increase in tryptophan solvent exposure. Studies on the environmental dependence of the tryptophan Fermi resonance couplet demonstrate that the 1360 cm^{-1} band is more intense in hydrophobic environments (Harada et al., 1986). This result, combined with the changes

in the 1430 cm^{-1} trytophan band, indicates that the tryptophans become more exposed to the solvent and better hydrogen bonded (presumably to the solvent) upon conversion to the molten globule state.

The amide III Raman band occurs at 1245 cm^{-1} and is more intense in molten globule BAL than it was in native BAL. It resembles very closely the amide III of denatured HEWL. Both the amide I and III Raman results appear to indicate that the molten globule state is quite close to that of the denatured state.

On the low frequency side of the amide III the tyrosine and phenylalanine vibration appears to be 5-10 cm^{-1} higher in frequency than it was in native BAL. The 1160 and 1180 cm^{-1} bands of native BAL appear to undergo some small shifts in frequency, which points to their involvement in backbone vibrations.

The ROA of this region is essentially a broad negative/positive couplet, peaking at 1225 and 1300 cm^{-1} and with almost no fine structure superimposed on it. It resembles the denatured HEWL amide III ROA very closely, the only difference being a small shift in the frequencies of the two couplets. This implication from this is the same as the interpretation of the Raman (and ROA) from the amide I and III regions; i.e. the molten globule state is quite close to the denatured state.

The most noticable change with the induction of the molten globule state is the disappearance of the positive 1340 cm^{-1} ROA band. This implies that the 1340 cm^{-1} band originates from a structural feature within the protein which disapppears as the protein moves away from the native state (or denatures). It is difficult to determine, purely from the BAL ROA spectra what this structural feature may be: even assigning this structure to tertiary or secondary structure is difficult because upon lowering the

pH native α-lactalbumin loses both these kinds of structure (Alexandrescu et al., 1993; Chyan et al., 1993). The assignment of the band is discussed further in chapter 8. Unfortunately, a denatured BAL ROA spectra proved difficult to obtain.

1150-700 cm^{-1}

The Raman bands in this region are similar to those found in the native BAL. The 1150 and 1080 cm^{-1} bands remain intact showing that they are insensitive to conformation. The ROA of the 1150 cm sidechain methyl rock remains positive (this region was prone to experimental artifacts), but the 1080 cm^{-1} aromatic C-H vibrations now gives a broad negative ROA feature. The latter vibration also changed more significantly upon HEWL denaturation and this suggests that the tryptophan vibrations are more conformationally sensitive, probably by mixing with backbone vibrations because the tryptophan ROA does not alter significantly between HEWL and denatured HEWL. The 1005 cm^{-1} phenylalanine band becomes more intense (as was observed by Yu). The positive ROA feature remains.

The C-C stretches between 930 and 950 cm^{-1} collapse showing intensity changes similar to those found by Yu. It has been proposed that these bands indicate α-helix (Wen et al., 1994a and Tu, 1986), and this observation concurs with the NMR conclusions about the molten globule state. However the experiments on Ca^{2+} free BAL show that the 930 cm^{-1} band is quite volatile. The Raman band at 880 cm^{-1} is the tryptophan band, it and the tyrosine bands at 850 and 830 cm^{-1} become clear in the molten globule state because of the buffer change. These three bands have the same relative intensities in BAL in water (Barron et al., 1992a). The ROA between 950 and 850 cm^{-1} appears to be positive, and this may indicate α-helix (Wen et al., 1994a), again agreeing with the conclusions of

NMR experiments. However, this positive region is in fact induced by baseline correction and so this conclusion must be treated with caution.

7.3.3 Apo, or metal free, BAL.

Three calcium free experiments were carried out: Ca^{2+} free BAL dissolved in a sodium borate buffer (Na^{+}-BAL), the previous sample but with $CaCl_2$ added (recalcified Na^{+}-BAL), and Ca^{2+} free BAL dissolved in Tris buffer (apo-BAL). All of these experiments are described together, but more of an emphasis is put on apo-BAL with the others being described only where the apo-BAL failled to give a result, or the results between the three experiments are different. The native BAL is refered to as Ca^{2+}-BAL.

The Raman and ROA spectra of Na^{+}-BAL, recalcalcified Na^{+}-BAL and apo-BAL are shown in Figures 7.9, 7.10 and 7.11, respectively.

1750-1500 cm^{-1}

The amide I Raman band appears at 1655 cm^{-1} is not dependent upon whether BAL is metal free or not. The amide I ROA is still discernible in the apo-BAL as a distinct couplet.

The aromatic and tryptophan vibrations appear to change intensity in the Ca^{2+} free experiments, but this is an experimental artifact and is caused by the decreasing solubilty of the Ca^{2+} forms of BAL. The ROA of the tryptophan vibrations remains the same in the Ca^{2+}-BAL, Na^{+}-BAL and recalcified Na^{+}-BAL.

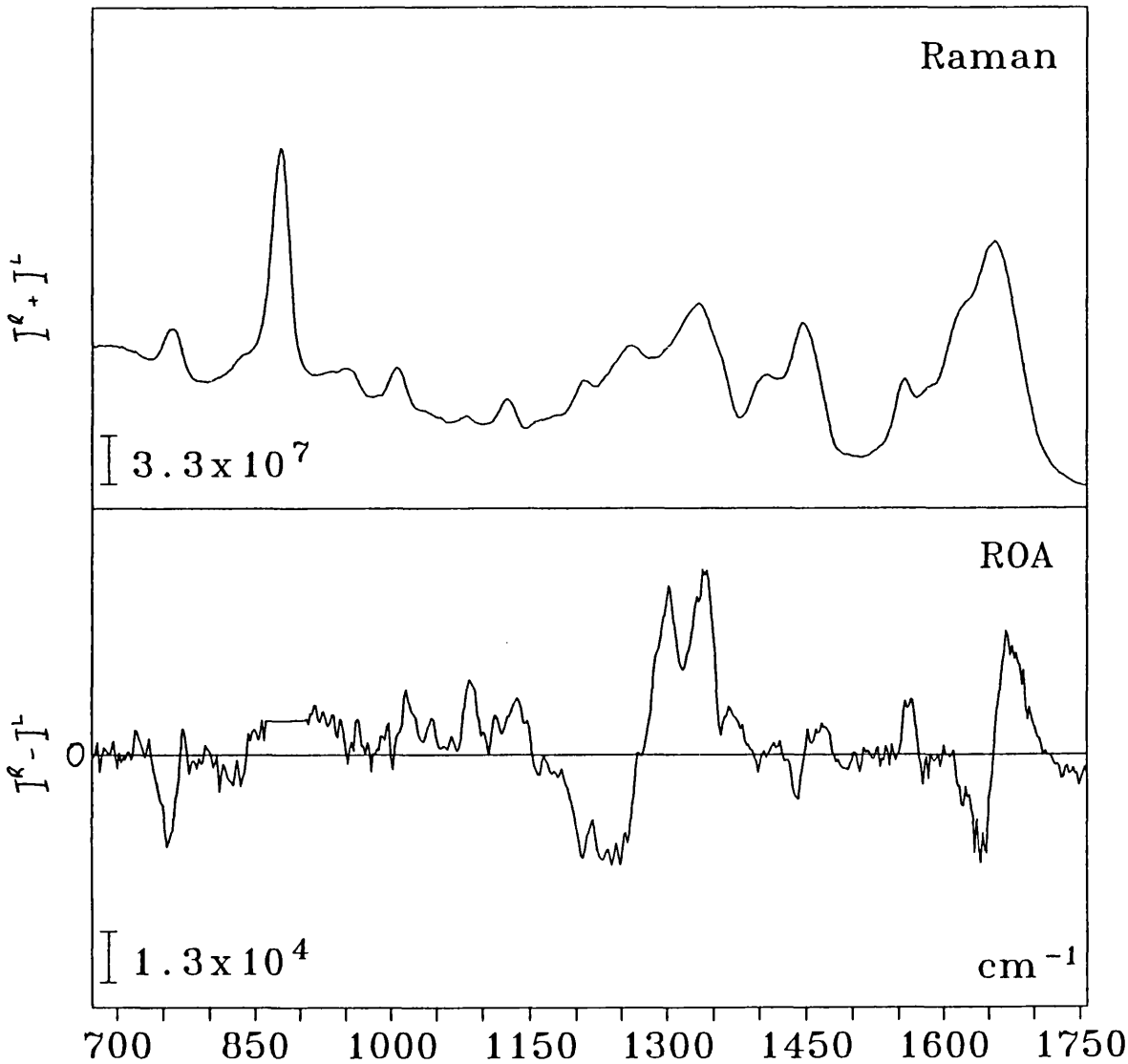


Figure 7.9: The Raman and ROA spectra of Na^+-BAL.

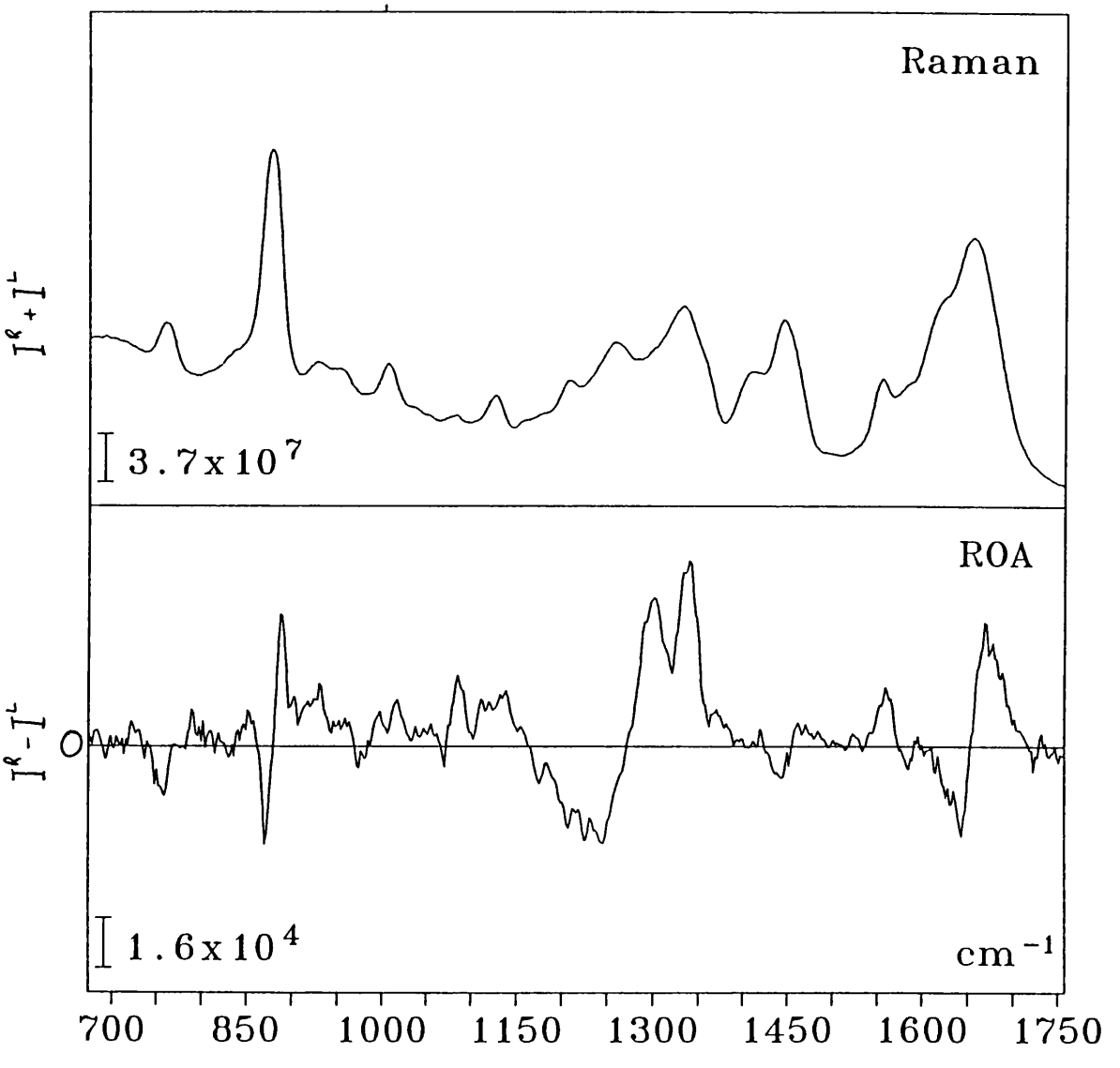


Figure 7.10: The Raman and ROA spectra of recalcified Na^+-BAL.

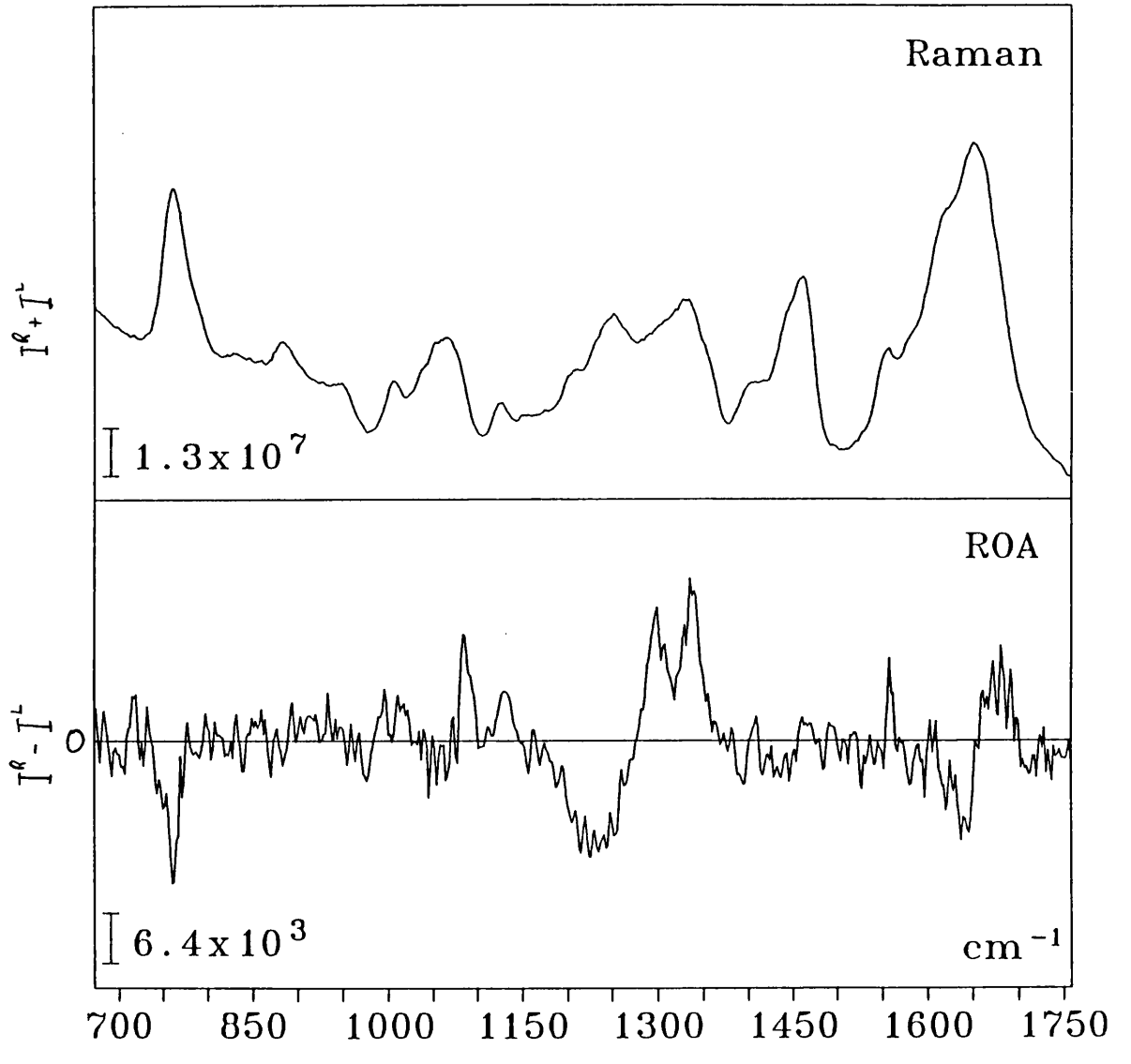


Figure 7.11: The Raman and ROA spectra of apo-BAL.

1500-1150 cm^{-1}

The CH_2 deformation Raman band appears to be more intense in the apo-BAL and increases in frequency, but this is due to a Tris buffer band at 1475 cm^{-1}. The Raman band is the same in Ca^{2+}-BAL, Na^{+}-BAL and recalcified Na^{+}-BAL. However, the ROA band in the latter two samples is a small negative/positive couplet and this differs from the negative peak of Ca^{2+}-BAL. (The ROA of apo-BAL is not clear enough to discern its behaviour in this region.) The reason for this difference is not known but a possible cause is the different methods of sample preparation (see section 7.2).

Although it appears as if the tryptophan Raman band at 1410 cm^{-1} is dependent on the metal binding, this is in fact not the case. All the 1410 cm^{-1} tryptophan Raman bands in Ca^{2+}-BAL, Na^{+}-BAL, recalcified Na^{+}-BAL and apo-BAL actually appear at the same frequency and the CH_2 deformations change. In apo-BAL the Tris buffer produces a large Raman band at ~1460 cm^{-1}.

The Raman bands between 1370 and 1150 cm^{-1} are the same in all the Ca^{2+} free samples as they were on Ca^{2+}-BAL. The only exception is the amide III band of the apo-BAL which seem to fall in frequency and increase intensity, becoming closer to the molten globule amide III Raman. Although this appears to contradict the ROA spectra, a closer examination reveals that at the low concentrations of apo-BAL used for this experiment significant buffer Raman band will be observed. This is shown in Figure 7.12.

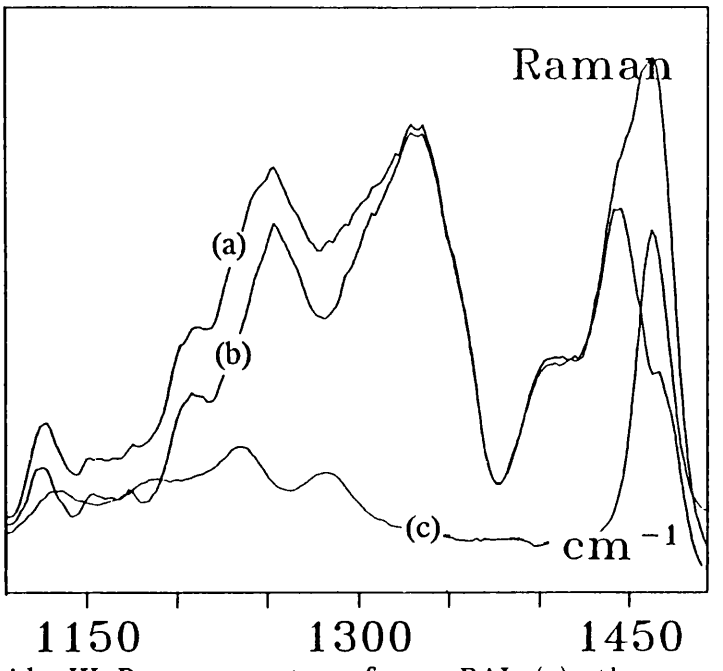


Figure 7.12: Amide III Raman spectra of apo-BAL (a), the apo-BAL spectra with Tris subtracted (b), and Tris alone (c).

The amide III ROA is almost exactly the same in apo-BAL as it is in Ca^{2+}-BAL. Particularly interesting is the positive 1340 cm^{-1} band which is very sensitive to structure - as was shown in the molten globule state - but remains completely intact in the apo-BAL. Another amide III band that may change is the positive/negative couplet between 1130 and 1190 cm^{-1} where the negative portion of the couplet appears to collapse upon decalcification. Although difficult to see in the apo-BAL spectrum, the same effect is seen in the Na^{+}-BAL and it's recalcified form. Upon recalcification of Na^{+}-BAL three changes to the ROA spectrum are observed; the regeneration of the negative 1170 cm^{-1} band as discussed above, the small increase in the 1340 cm^{-1} band, and the slightly better definition of the 1225 and 1245 cm^{-1} bands. If such changes are real they are small, and they will only be proved to be real by the reproducibility of the effects. Differences do appear between the recalcified Na^{+}-BAL and Ca^{2+}-BAL, even although they should ideally be the same sample. One noticable difference is the extra negative band at 1200 cm^{-1}, which must

be an experimental artifact because it is present in the Na^+-BAL too. The other differences between recalcified Na^+-BAL and Ca^{2+}-BAL, such as the extra intensity in the 1340 cm^{-1} band. Some of the differences could originate from an to the excess of calcium present in the recalcified Na^+-BAL (x20) compared to the Ca^{2+}-BAL (x2).

In no way does the amide III ROA indicate such a drastic change as occurs in the native to molten globule transition, a change which other experimental techniques, such as fluorescence, do indicate (Permyakov et al., 1981). The difference between ROA and techniques like fluorescence is that the latter makes the assumption that changes in the sidechain structure, or environment, must be reflected, or caused by, changes in the backbone structure. ROA does not have to make such an assumption and techniques which examine a protein molecule as a whole, such as ROA and NMR, agree that the differences between apo- and Ca^{2+}-BAL are small. The positive 1340 cm^{-1} ROA band remains in place, one possible assignment for this band is surface loops including the Ca^{2+} binding loop (see chapter 8.)

The fluorescence spectra below in Figure 7.13 show clearly that the apo-BAL has the characteristic fluorescence of the metal free form of the protein and this is further proved by the hypochromic wavelength shift and the quenching that occurs when the BAL binds to metal ions. The apo-BAL samples that were used for the ROA experiment showed similar fluorescence behaviour when examined after the ROA experiment. The binding of Na^+ ions is dependent upon concentration, but this data concurs with the Na^+ binding constant of ~100 M^{-1} (Hiraoka and Sugai, 1985; McKenzie and White, 1991; Desmet et al., 1987).

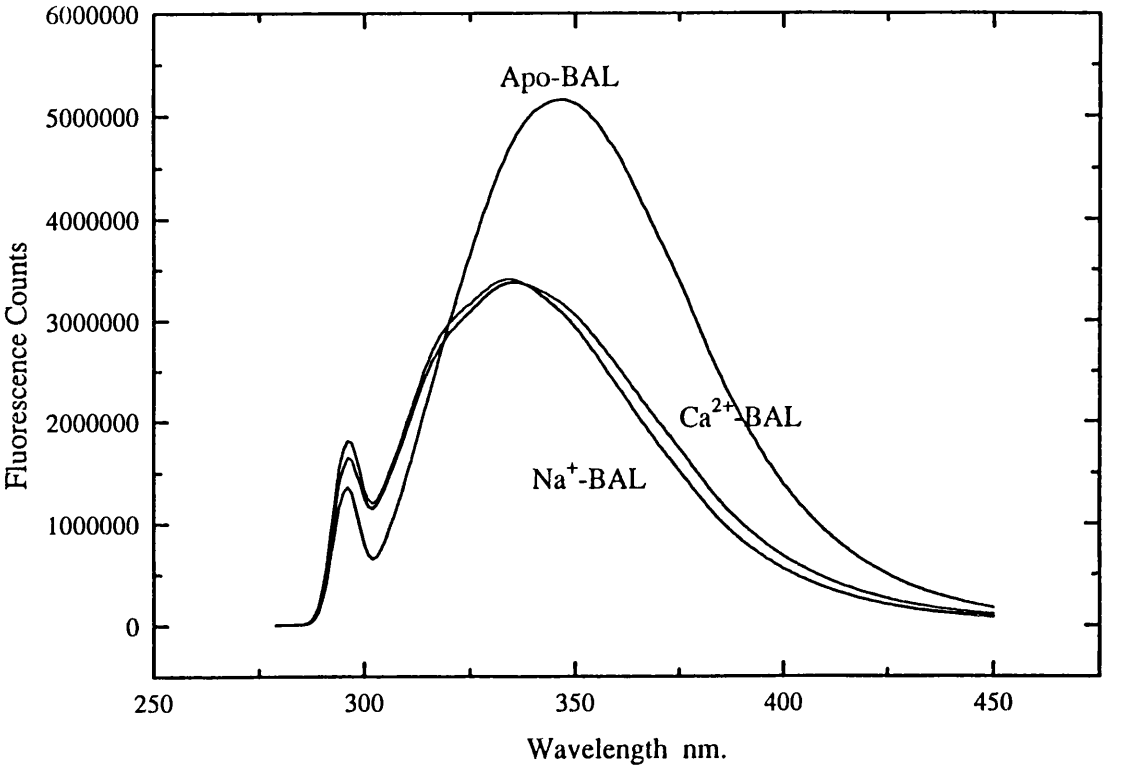


Figure 7.13: Fluorescence spectra of apo-BAL, Na^+-BAL and Ca^{2+}-BAL. (Baseline subtracted and concentration normalized.)

The fluorescence experiments were run in 100mM Tris buffer. UVCD experiments by Hiraoka and Sugai in varying concentrations of Tris buffer have demonstrated that Tris can stabilize the holo-BAL form (Hiraoka and Sugai, 1985).

1150-700 cm^{-1}

The Raman bands of the apo-BAL are hard to determine. Tris buffer bands appear at 1056 and 782 cm^{-1}. The main differences between the Ca^{2+}-BAL and apo-BAL are that the 900 cm^{-1} C-C stretch and 830 and 850 cm^{-1} tyrosine vibrations seen to disappear, or be swallowed up in the buffer baseline.

The Raman bands in this region in the Na^+-BAL, its recalcified form and Ca^{2+}-BAL remain very similar to each other. The only exception is the C-C stretch band where the intensity of the 930 cm band increases in the sequence Ca^{2+}-BAL, Na^+-BAL and recalcified Na^+-BAL. Perhaps this band envelope is also affected by the experimental procedures in the same way the 1400-1460 cm^{-1} band envelope was changed. The ROA of the Ca^{2+}-BAL, Na^+-BAL and recalcified Na^+-BAL are almost identical. The ROA of the apo-BAL is not clear.

Chapter 8

Amide ROA signatures of proteins.

8.1 THE AMIDE I REGION

Conventional vibrational spectroscopy has made extensive use of the amide I and I' vibrations with several high resolution FTIR spectra of various proteins, and the subsequent structural predictions, being published recently (Prestrelski et al., 1991a, b, c and references contained therein). However, recent normal mode calculations throw some doubt on the value of the conventional amide I Raman and IR frequency-structure correlations. Krimm and Reisdorf (1994) have concluded that the small geometric irregularities and force constant variations in α-helical structures can alter the amide I frequencies by significant amounts. It has been proposed that since VOA techniques provide another criteria by which protein structure can be determined, the frequency variations will not have a significant effects on VOA measurements. VCD has used the amide I' region for structural prediction with reasonable success (Keiderling et al., 1994; Pancoska et al., 1994).

Structural determination from the amide I ROA is more difficult. There are three main reasons for this: first the amide I ROA is experimentally difficult to obtain accurately, second the amide I does not span a broad range of frequencies (and so all the information is crammed into a small space), and third the amide I ROA from different structures can be very similar (see Fig. 3 in Barron et al., 1994). The positively biased amide I couplet found in denatured HEWL is also seen in other folded proteins such as BSA, ribonuclease (Barron et al., 1992) and insulin (G. Wilson, unpublished data).

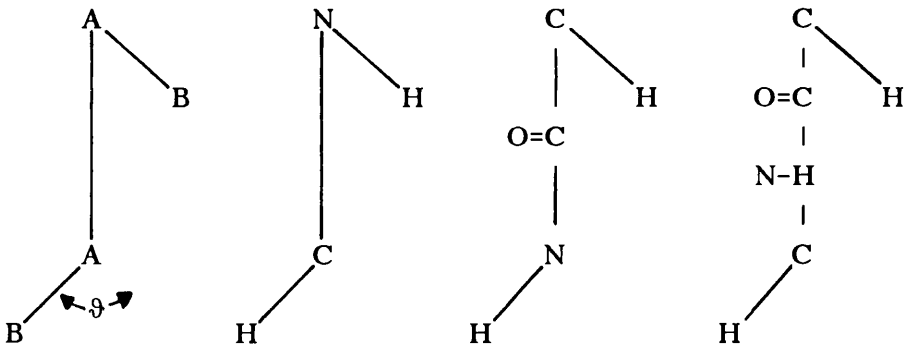
8.2 THE ROA SIGNATURES OF THE AMIDE III REGION

8.2.1 Background

The amide III region will probably become the most significant part of the protein ROA spectrum, and it will do so for several reasons. Firstly, the amide III ROA appears very quickly during a protein ROA acquisition, and can be see with poor quality samples (as in the apo-BAL). Secondly, it is one of the most reliable regions of the ROA spectrum, unlike other regions such as amide I, because the amide III vibrations are non-polarized. Thirdly, it appears to differ significantly between different protein samples, containing the most noticable changes between two ROA spectra.

8.2.2 Theoretical aspects

The amide III mode is made up of C-H and N-H deformations (Diem et al., 1992; Ford et al., 1994) which, as they are vibrations of particular groups that are separated by another bond, can be treated like a two group model. The 'traditional' two-group model considers two neutral equivalent groups attached to each other which contributed equally to the resulting normal modes generated by the overall structure; such a structure is shown in Figure 8.1(a). If the amide link is considered as Figures 6.4.1(b), (c) and (d) then it can be seen how it fits into the two group model. Although far from the ideal two group model it can be seen that theoretical conclusions from the two group model can be useful in describing some of the amide III spectral features.



(a) (b) (c) (d)

Figure 8.1: The two group model (a), the three alternative ways of considering the amide link as a two group structure (b), (c) and (d).

The theoretical ROA and VCD characteristics of various coupled vibrations have been deduced previously (Barron, 1982). An extention of the previous 90° scattering equations for the two group model (Barron, 1982) gives a backscattering equivalent (Barron et al., 1994):

$$\Delta^{\pm}(180^{\circ}) = \mp 8\pi R_{21}\sin\vartheta/7\lambda(1\mp\cos\vartheta) = K(\mp\sin\vartheta/1\mp\cos\vartheta)$$

However, the sign of the coupled deformations is given by the simplified equation:

$$\Delta^{\pm}(180^{\circ}) = \mp \sin\vartheta \text{ (as long as } \vartheta \neq 0^{\circ} \text{ or } 180^{\circ})$$

Although the two group model assumes that the two groups are axially symmetric and the connecting bond is rigid, further investigations do prove interesting. In order to establish the relationship between this theoretical two group equation and protein conformation we must determine how the torsion angle ϑ can be deduced from the Φ and Ψ values. For the peptide

defined in Figure 8.1 ϑ is related to Φ and Ψ by the following relationships (which assume a planar trans amide link):

C_n-H to N_n-H;	$\vartheta = \Phi - 60^o$	Φ dependence
N_n-H to C_{n+1}-H;	$\vartheta \approx \Psi - 120^o$	Ψ dependence
C_n-H to C_{n+1}-H;	$\vartheta \approx \Phi + \Psi + 180^o$	Combined Ψ/Φ dependence

(Note that the Φ torsional angles are defined as those across the bond C(=O)-C_α-N-C(=O) and Ψ as N-C(O)-C_α-N, a positive rotation is defined as a clockwise rotation of the rear bond as viewed down the bond axis.)

Can the two group model and the peptide torsion angles explain the amide III ROA? For the ab initio calculations, the answer is yes. In their examination of alanine dipeptide Polavarapu and Deng (1994) have produced a series of amide III ROA spectra for various conformations. Their results are tabulated in Table 8.1, along with the signs of the CID for the antisymmetric/symmetric couplet, as derived from the two group model for the models of deformation coupling dependence on Φ, Ψ and the combined Ψ/Φ angles. (There is no theoretical basis for the assignment of the low and high frequency components of the couplet to the antisymmetric and symmetric combinations, respectively. However, as can be seen from Table 8.1, the ab initio and two group model results fit better when we make this assumption.

It can be seen form Table 8.1 that only the Φ dependence model correctly predicts all four ab initio amide III results. Although the Ψ dependence model correctly predicts two results, it fails to predict the signs found in the ab initio results for conformers with significantly different Φ angles. Although the Φ dependence model is based on an

Table 8.1: A comparison (using four conformers) of the amide III couplets signs from ab initio ROA calculations and the various simple two group models proposed above. The correct two group model couplet signs are shown in bold.

Conformational angles		Φ dependence model		Ψ dependence model		Combined Ψ/Φ dependent model		Ab initio amide III sign
Φ(°)	Ψ(°)	ϑ(°)	Sign	ϑ(°)	Sign	ϑ(°)	Sign	
-160	-40	+140	**+/-**	-160	**-/+**	-20	-/+	+/-
-60	-40	-120	**-/+**	-160	**-/+**	80	+/-	-/+
-100	120	-160	**-/+**	0	**0**	-160	**-/+**	-/+
-150	160	150	**+/-**	40	**+/-**	-170	-/+	+/-

equation from a very simplistic coupling model it has yielded some interesting correlations with the more complex ab initio ROA predictions.

The next step would be the correlation of the Φ dependence model with experimentally determined ROA. A negative/positive amide III couplet can be explained by the Φ-dependence model. In proteins the majority of the amide Φ angles are between -120° and 60° (see Figure 8.2), which is in the negative/positive amide III range according to the Φ-dependence model.

The amide III vibration would be expected to be more sensitive to Φ angles than Ψ angles, because in the latter the vibration must mix across the carbonyl group. This conclusion is supported by other data: Krimm and Bandekar (1986) found an amide III dependence on Φ in their normal mode calculations, and a possible Φ-dependence was observed in the alanyl peptides (see section 5.5.2). Although, this simple model explains ab initio and experimental amide III ROA results, the argument has one flaw: the ab initio alanine dipeptide data points to a Φ-dependence model, the experimental alanine dipeptide data does not (see section 5.3.3). Although

there are reasonable explanations as to why the ab initio and experimental data do not match (and in fact the alanine dipeptide Φ angle remains in the -120° to 60° range in both H_2O and $CHCl_3$), should this idea be expanded this discrepance will have to be rationalized.

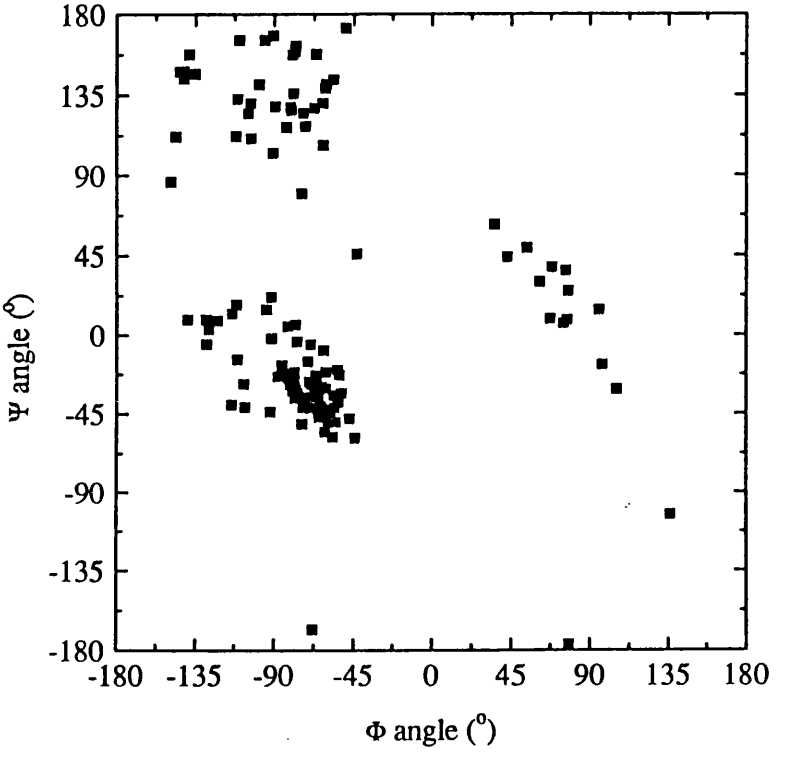


Figure 8.2: The Ramanchandran plot for HEWL.

8.2.3 Earlier work

Earlier ROA spectra of proteins and peptides were recently published by this laboratory (Wen et al., 1994a and b). As was pointed out in these papers many other studies would have to be undertaken before the ROA-protein correlations derived in that paper could be applied with confidence. Therefore, it is important to understand these assignments and discuss what effect these earlier results have on the assignment of the HEWL and α-lactalbumin spectra, and vice versa.

The model peptide used to simulate α-helix was that of poly-L-lysine (PLL). Traditionally used as a model peptide, PLL has the advantage that it

can adopt all three major secondary structure groups (random coil, α-helix, β sheet). PLL was used to obtain ROA spectra of random coil and α-helix.

However, in recent years it has become apparent that the structures formed by PLL might deviate from the ideal secondary structure. UV circular dichroism and vibrational circular dichroism (Tiffany and Krimm, 1968; Paterlini et al., 1986) have suggested that the PLL 'random coil' may in fact be an extended helical structure. This extended helix arises because the Ψ and Φ angles that any polypeptide can adopt are restricted. In the case of PLL this restriction will be enhanced because all the sidechains are the same, and the lysine sidechains are so large. What influence this semi-regularity will have on the ROA spectra is not clear and the problem is compounded because there are no shorter lysine peptide ROA spectra to draw comparisons with; any delocalization, or broadening, of the amide III band cannot be tested.

The α-helix structure has also been questioned in the last few years and anomalous α-helix formation in lysine peptides, particularly with long chain lysine peptides, has been reported (Jackson et al., 1989). It was proposed that the α-helical structure co-exists with either random coil, β sheet, turns or alternative hydrogen bonding.

In the random coil to α-helix transition of PLL there were three regions of the ROA spectrum which change and are ascribed to α-helix by Wen et al. (1992a). They are a broad, positive band between 915 and 990 cm^{-1}, a negative/positive couplet at 1103 cm^{-1} and a negative/positive couplet at 1275 cm^{-1}.

The first two α-helix features are seen in both insulin and bovine serum albumin (BSA). The third is difficult to distinguish in proteins where there is a large couplet crossing over at this point anyway. (A large

negative/positive couplet can be seen, at similar frequencies, in the ROA spectrum of denatured HEWL.)

A fourth change also occurred in the amide III region between random coil and α-helix, this was a strong positive 1340 cm^{-1} band. The amide III region of α-helix PLL is very similar to the amide III region of BSA and both show the strong, positive band at 1340 cm^{-1}. This band is also found in HEWL where it was shown to be an amide III band. Its absence in insulin indicates that it cannot be due to α-helix. This 1340 cm^{-1} band is found in the amide III region of BSA, α-lactalbumin and HEWL. The α-helix PLL experiment has recently been repeated under stricter temperature control (G. Wilson et al., manuscript in preparation), which showed that the amide III region remained the same.

The origins of this 1340 cm^{-1} signal is very probably the 'other' structure that has been observed in α-helical PLL. It is not unreasonable, but perhaps speculative, to postulate that α-helical sections are attracted towards each other due to hydrophobicity (the lysine is neutral charged at high pH and has a four CH_2 groups on it's sidechain). This would generate a helix-loop-helix structure.

Another hypothesis is that the 'other' structure is a region of random coil, or a loop, maintained by disulphide bonds. Such loops are found in BSA, α-lactalbumin and HEWL. However, that cannot explain the presence of the 'other' structure in PLL or its absence in molten globule α-lactalbumin and denatured HEWL.

The BSA results lead to the conclusion that the band was due to rigid helix-loop-helix conformations - the rigidity being conferred by disulphide bridges. The structure of human serum albumin is known and it has 17 disulphide bridges - as does BSA - and 28 sections of α-helix (He and

Carter, 1992; Peters, 1985). However, this implies that the 'other' structures in 'α-helical' PLL are rigid loops between the helical sections. Rigidity of the PLL loops may be confered by the large size of the lysine residues which will restrict the conformational space available during dymanic motion.

8.2.4 Assignment of the 1340 cm^{-1} band.

One of the originally suggested ideas was that the 1340 cm^{-1} band was due to rigid loop structures, or rigid helix-loop-helix structures, and this hypothesis fits with most of the experimental evidence. However, in order to fit this idea into the HEWL and BAL results we must find loop structures within these two proteins that are more rigid in BAL than HEWL, and will become more rigid in HEWL upon NAG_3 binding. Although it is conventionally accepted that α-lactalbumin has more dynamic flexiblity than HEWL a comparison of the main chain temperature B-values of α-lactalbumin and HEWL indicate that the surface loop HEWL residues 69-72 (which alters during NAG_3 binding - see section 6.1.5) have substantially higher disorder than their α-lactalbumin counterparts. This loop region around 65 to 75 in HEWL appears to be an ideal candidate for the rigid loop hypothesis. The relationship between ROA signal and the dynamic nature of a polypeptide could be examined by examining protein systems where a small chemical change induces a large dynamic change; i.e. trypsin and trypsinogen. One disadvantage of this idea is that there is no capacity for it within the present theoretical framework. (The nature of dynamically related ROA may depend upon the quantum mechanical tunnelling through conformational barriers between subtly different conformational states (Gol'danskii et al., 1989).) The peptide ROA results indicated that some of the difference between Val_2 and Ala_2, may lie with different molecular dynamics.

Although the rigid loop hypothesis fits this data, it is not possible to assign the 1340 cm^{-1} band to the 65-75 region specifically. The NMR HEWL hydrogen exchange data indicates that most of the residues between 65 to 75 are category I amides and exchange very quickly. Looking at BSA, which contains the strongest 1340 cm^{-1} signal, we see that the band vanishes quickly upon deuteration (Barron et al., 1995). However, in HEWL this is not the case and the 1340 cm^{-1} band remains relatively intense until complete deuteration, indicating that, although deuterium sensitive, the band originates from slow exchange amides. This may be due to other areas of the HEWL structure that are rigid and slow exchanging. (Although there is no correlation between hydrogen exchange rates and temperature factors, local fluctuations do play a significant role in amide hydrogen exchange (Radford etal., 1992).)

In α-lactalbumin the 1340 cm^{-1} band in BAL is similar to that in apo-BAL. The Ca^{2+} binding loop (79-88) in α-lactalbumin does not correspond to the 65-70 loop region in HEWL (see Figure 7.1). However, both these BAL and HEWL loops come within the same 20 residue two disulphide double loop region between 64 and 94 in HEWL and 61 to 91 in BAL (see Figure 7.1). The rigid loop assignment of the 1340 cm^{-1} band indicates that the Ca^{2+} binding loop of BAL is not significantly altered in the apo-BAL used in these ROA experiments.

8.2.5 Assignment of the 1300 to 1190 cm^{-1} bands.

The positive 1300 cm^{-1} ROA signal has contributions from conformationally and deuteration insensitive vibrations. In HEWL a significant part of the intensity of this band disappears only when the protein is completely deuterated. In BSA the 1300 cm^{-1} band appears to remain invariant upon partially deuteration (Barron et al., 1994). All this information concurs with the assignment of Wen et al. (1992a) of this band

to α-helix. However, care will have to be taken with this assignment because overall the band has contributions from non-helical structure and possibly C-H vibrations.

The negative 1270 cm^{-1} negative band originates from conformationally sensitive C-H deformations which have a very small N-H dependence. The structure that causes this feature is specific to HEWL.

The negative 1245 cm^{-1} band in BAL is also seen in denatured HEWL at 1235 cm^{-1}. This feature is very prominant in the ROA spectrum of insulin and was called the 'insulin type loop' (Wen et al., 1994a) - although 'insulin type structure' may be more appropriate. It is also found in ribonuclease (Barron et al., 1992), concanavalin A and α-chymotrysin (Wen et al., 1994b). The 1245 cm^{-1} band is deuterium sensitive and so is not due to C-H deformations. It is of a similar intensity in denatured HEWL as in the native HEWL and vanishes very quickly upon deuteration. It has been proposed that this band originates from dynamically fluctuating structure, such as surface loops or end chains. Such mobile structures would be found in a significant number of proteins. This hypothesis could only be tested by studying a system where rigid and fluctuating states could be induced, or by establishing that this band could not be due to any specific structure. However, another possiblity is that this band may be due to an amide III vibration from amide links which occupy a very common low energy position on the Ramachandran map, subsequently many proteins and denatured proteins will have this 1245 cm^{-1} feature. The latter possibility may explain why this band has been observed in low temperature HEWL ROA studies (G. Wilson,manuscript in preparation).

The negative 1220 cm^{-1} ROA signal in native BAL appears to originate from a BAL specific structure, but a significant part of this band remains in the molten globule state of BAL. The band cannot be due to α-helix

(which is found in the molten globule state of BAL) or other common structural or dynamic effects because it is specific to BAL. It's assignment is uncertain.

The 1130 to 1240 cm^{-1} positive/negative couplet appears to be related to α-helix. Although this couplet is not particularly prominent in HEWL or BAL, it is large in insulin and BSA, although the latter ROA spectrum is dominated by the 1340 cm^{-1} band (Barron et al., 1994). The positive component of the couplet is very small in proteins with low α-helical content such as chymotrypsin and ribonuclease. The negative component seems to be equally present in all protein samples. The deuterated HEWL spectra indicates that the negative component of this couplet is easily exchanged and probably has separate origins to the positive component.

8.3 SUMMARY

The amide ROA assignments that can be concluded from, or are supported by, the work in this section of the thesis are given in Table 8.2.

Table 8.2: Summary of the protein structure/conformation ROA assignments concluded from this work.

Vibration or ROA band (cm^{-1})	Assignment	Comments
Amide I	-	A positively biased couplet may indicate α-helix, but most protein structural features appear to give negative/positive couplets.
1340 cm^{-1}	Rigid loops	Although not proved, there is compelling evidence for this assignment.
1300 cm^{-1}	-	Possibly α-helix.
1245 cm^{-1}	Dynamic structure	This band appears to originate from dynamic or disordered protein structure (such as surface loops and end chains).
1220 cm^{-1}	-	Possibly disordered structure.
1150 cm^{-1}	α-helix	

Acharya, K.R., Stuart, D.I., Walker, N.P.C., Lewis, M., Phillips, D.C., (1989), J. Mol. Biol., **208**, 99-127

Acharya, K.R., Ren, J., Stuart, D.I., Phillips, D.C., (1991), J. Mol. Biol., **227**, 571-581

Alexandrescu, A.T., Broadhurst, R.W., Wormald, C., Chyan, C.L., Baum, J.,Dobson, C.M., (1992), Eur. J. Biochem., **210**, 699-709

Alexandrescu, A.T., Evans, P.A., Pitkeathly, M., Baum, J., Dobson, C.M., (1993), Biochemistry, **32**, 1707-1718

Alexandrescu, A.T., Ng, Y.L., Dobson, C.M., (1994), J. Mol. Biol., **235**, 587-599

Alsenoy, C.V., Cao, M., Newton, S.Q., Teppen, B., Perczel, A., Csizmadia, I.G., Momany, F.A., Schafer, L., (1993), Journal of Molecular Structure (Theochem.), **286**, 149-163

Amos, R.D., (1982), Chem. Phys. Lett, **87**, 23-26

Aramini, J.M., Drakenberg, T., Hiraoki, T., Ke, Y., Nitta, K., Vogel, H.J., (1992), Biochemistry, **31**, 6761-6768

Avignon, M., Garrigou–lagrange, C., Bothorel, P., (1973), Biopolymers, **12**, 1651-1669

Bandekar, J., (1992), Biochem. Biophys. Acta, **1120**, 123-143

Bandekar, J., Krimm., S., (1988), Biopolymers, **27**, 885-908

Barron, L.D., (1982), Molecular Light Scattering and Optical Activity, Cambridge University Press, Cambridge,

Barron, L.D., Gargaro, A.R., Hecht, L., Polavarapu, P.L., (1991), Spectrochim. Acta, **47A**, 1001-1016

Barron, L.D., Cooper, A., Ford, S.J., Hecht, L., Wen, Z.Q., (1992a), Faraday Discuss., **93**, 259-268

Barron, L.D., Gargaro, A.R., Hecht, L., Polavarapu, P.L., (1992b), Spectrochim. Acta, **48A**, 261

Barron, L.D., Gargaro, A.R., Hecht, L., Polavarapu, P.L., Sugeta, H., (1992c), Spectrochim. Acta, **48A**, 1051-1066

Barron, L.D., Ford, S.J., Bell, A.F., Wilson, G., Hecht, L., Cooper, A., (1994), Faraday Discuss., **99**, in press

Baum, J., Dobson, C.M., Evans, P.A., Hanley, C., (1989), Biochemistry, **28**, 7-13

Berliner, L.J., Koga, K., Nishikawa, H., Scheffler, J.E., (1987), Biochemistry, **26**, 5769-5774

Bernard, M., Canioni, P., Cozzone, P., Berthou, J., Jolles, P., (1990), Int. J. Peptide. Protein. Chem., **36**, 46-55

Bertoluzza, A., Bonora, S., Fini, G., Morelli, M.A., (1992), Canadian J. Appl. Spec., **37**, 58-61

Birke, S.S., Agbaje, I., Diem, M., (1992), Biochemistry, **31**, 450-455

Blake, C.C.F., Cassels, R., Dobson, C.M., Poulsen, F.M., Williams, R.J.P., Wilson, K.S., (1981), J. Mol. Biol., **147**, 73-95

Brooks, C.L., Karplus, M., Pettitt, B.M., (1988), Advances in Chemical Physics, **71**

Buck, M., Radford, S.E., Dobson, C.M., (1993), Biochemistry, **32**, 669-678

Cassels, R., Dobson, C.M., Poulsen, F.M., Williams, R.J.P., (1978), Eur. J. Biochem., **92**, 81-97

Cheam, T.C., (1993), J. Mol. Structure, **295**, 259-271

Cheetham, J.C., Artymink, P.J., Phillips, D.C., (1992), J. Mol. Biol., **224**, 613-628

Chen, M.C., Lord, R.C., Mendelsohn, R., (1973), Biochem. Biophys. Acta, **328**, 252-260

Chen, M.C., Lord, R.C., Mendelsohn, R., (1974), J. Am. Chem. Soc., **96**, 3034-3042

Chyan, C.L., Wormald, C., Dobson, C.M., Evans, P.A., Baum, J., (1993), Biochemistry, **32**, 5681-5691

Cooper, A., (1974), Biochemistry, **13**, 2853-2856

Cooper, A., (1980), Sci. Prog., Oxf., **66**, 473-497

Cooper, A., Eyles, S.J., Radford, S.E., Dobson, C.M., (1992a), J. Mol. Biol., **225**, 930-943

Cooper, A., Johnson, C.M., (1994), in Methods in Molecular Biology: Physical methods of analysis, **22**, C. Jones, B. Mulloy, A.H. Thomas (Eds), Humana Press, Clifton, N.J., Chapters 9 and 10

Covington, A.K., Paabo, M., Robinson, R.A., Bates, R.G., (1968), Analytical Chem., **40**, 700-706

Creighton, T.E., (1984), Proteins, Structure and Molecular Properties, Freeman, NY, 159-198 (Chapter 5)

Creighton, T.E., (1989), in Protein Structure: A Practical Approach, T.E. Creighton (Ed), IRL Press, Oxford, 155-167 (Chapter 7)

Delepierre, M., Dobson, C.M., Karplus, M., Poulsen, F.M., States, D.J., Wedin, R.E., (1987), J. Mol. Biol., **197**, 111-130

Desmet, J., Hanssens, I., van Cauwelaert, F., (1987), Biochem. Biophy. Acta, **912**, 211-219

Diem, M., Lee, O., Roberts, G.M., (1992), J. Phys. Chem., **96**, 548-554

Evans, P.A., Topping, K.D., Woolfson, D.N., Dobson, C.M., (1991), Protein: Structure, Function and Genetics, **9**, 248-266

Fasman, G.D., Itoh, K., Liu, C.S., Lord, R.C., (1978), Biopolymers, **17**, 125-143

Ford, S.J., Wen, Z.Q., Hecht, L., Barron, L.D., (1994), Biopolymers, **34**, 303-313

Fukamizo, T., Ikeda, Y., Ohkawa, T., Goto, S., (1992), Eur. J. Biochem., **210**, 351-357

Gargaro, A.R., (1991), Doctoral Thesis, Glasgow University

Gargaro, A.R., Barron, L.D., Hecht, L., (1993), Journal of Raman Spectroscopy, **24**, 91-96

Gill, S.C., Von Hippel, P.H., (1989), Analytical Biochem., **182**, 319-326

Gould, I.R., Hillier, I.H., (1993), J. Chem. Soc., Chem. Commun., 951-952

Haezebrouck, P., Van Dael, H., (1993), Journal of Molecular Structure, **294**, 143-146

Han, S.J., Kang, Y.K., (1993), Int. J. Peptide Protein. Res., **42**, 518-526

Harada, I., Takeuchi, H., (1986), in Spectoscopy of Biological Systems, Advances in Spectroscopy, **13**, R.J.H. Clark and R.E.Hester (Eds), Wiley, Chichester, 113-175 (Chapter 3)

Harada, I., Miura, T., Takeuchi, H., (1986), Spectrochim. Acta, **42A**, 307-312

He, X.M., Carter, D.C., (1992), Nature, **385**, 209-215

Hecht, L., Barron, L.D., (1990), Applied Spec., **44**, 483-491

Hecht, L., Barron, L.D., Gargaro, A.R., Wen, Z.Q., Hug, W., (1992), Journal Raman Spec., **23**, 401-411

Hecht, L., Nafie, L.A., (1991), Molecular Physics, **72**, 441-469

Helgaker, T., Ruud, K., Bak, K.L., Jørgensen, P., Olsen, J., (1994), Faraday Discuss., **99**, in press

Heremans, K., Wong, P.T.T., (1985), Chem. Phys. Lett., **118**, 101-104

Himmler, H.J., Eysel, H.H., (1989), Spectrochim. Acta, **45A**, 1077-1083

Hiraoka, Y., Sugai, S., (1985), Int. J. Peptide Protein Res., **26**, 252-261

Imoto, T., Johnson, L.N., North, A.C.T., Phillips, D.C., Rupley, J.A., (1972), in The Enzymes (3th edition), **7**, P.D.Boyer (Ed), Academic Press, New York, 665-868

Jackson, M., Haris, P.I., (1989), Chapman, D., Biochim. Biophys. Acta, **998**, 75-79

Keiderling, T.A., (1994), Faraday Discuss., **99**, in press

Keiderling, T.A., Wang, S., Urbanova, M., Pancoska, P., Dukor, R.K., (1994), Faraday Discuss., **96**, in press

Koca, J., Carlsen, P.H.J., (1993), Journal of Molecular Structure, **291**, 271-286

Krimm, S., Bandekar, J., (1986, Adv. in Prot. Chem., **38**, 181-354

Krimm, S., Reisdorf, W.C., (1994), Faraday Discuss., **99**, in press

Kumagai, I., Maenaka, K., Sunada, F., Takeda, S., Miura, K., (1993), Eur. J. Biochem., **212**, 151-156

Kuwajima, K., Harushima, Y., Sugai, S., (1986), Int. J. Peptide Protein Res., **27**, 18-27

Lee, O., Roberts, G.M., Diem, M., (1989), Biopolymers, **28**, 1759-1770

Lord, R.C., (1977), Appl. Spectroscopy., **31**, 187-194.

Lord, R.C., Yu, N.T., (1970), J. Mol. Biol., **50**, 509-524

Lumb, K.J., Cheetham, J.C., Dobson, C.M., (1994), J. Mol. Biol., **235**, 1072-1089

McKenzie, F.A., White, F.H., (1991), Adv. in Prot. Chem., **41**, 124-316

McCammon, J.A., Harvey, S.C., (1987), Dynamics of proteins and nucleic acids, Cambridge University Press, Cambridge

Machida, K., Izumi, M., Kagayama, A., (1979), Spectrochim. Acta., **35A**, 1333-1339

Madison, V., Kopple, K.D., J. Am. Chem. Soc., (1980), 102, pp4855-4863

Malon, P., (1994), Faraday Discuss., **99**, in press

Milner-White, E.J., Poet, R., (1987), Trends in Biological Science, **12**, 189-192

Milranker, A., Radford, S.E., Karplus, M., Dobson., C.M., (1991), Nature, **349**, 633-636

Miura, T., Takenchi, H., Harada, I., (1989), J. Raman Spec., **20**, 667-671

Miura, T., Takenchi, H., Harada, I., (1991), Biochemistry, **30**, 6074-6080

Nafie, L.A., (1991), 4th Int. Conference. on C.D. (Bochum), 101-114

Naik, V.M., (1992), Vibrational Spectroscopy, **3**, 105-113

Oboodi, M.R., Alva, C., Diem, M., (1984), J. Phys. Chem., **88**, 501-505.

Pancoska, P., Yasui, S.C., Keiderling, T.A., (1989), Biochemistry, **30**, 5089-5103

Pancoska, P., Biffo, E., Janota, V., Keiderling, T.A., (1994), Faraday Discuss., **99**, in press

Paterlini, M.G., Freedman, T.B., Nafie, L.A., (1986), Biopolymers, **25**, 1751-1765

Pederson, T.G., Sigurskjold, B.W., Andersen, K.V., Kjaer, M., Poulsen, F.M., Dobson, C.M., Redfield, C., (1991), J. Mol. Biol., **218**, 413-426

Permyakov, E.A., Yarmolenko, V.V., Kalinichenko, L.P., Morozoa, L.A., Burstein, E.A., (1981), Biochem. Biophy. Res. Comm., **100,** 191-197

Peters, T.Jr., (1985), Adv. in Prot. Chem., **37**, 161-245

Polavarpu, P.L., (1990), J. Phys. Chem., **94**, 8106

Polavarapu, P.L., Deng, Z., Ewig, C.S., (1994), J. Am. Chem. Soc., **98**, 9919-9930

Polavarapu, P.L., Deng, Z., (1994), Faraday Discuss., **99**, in press

Prestrelski, S.J., Byler, M.D., Thompson, M.P., (1991a), Biochemistry, **30**, 8797-8804

Prestrelski, S.J., Byler, M.D., Thompson, M.P., (1991b), Int. J. Peptide Protein Res., **30**, 8804-8811

Prestrelski, S.J., Byler, M.D., Thompson, M.P., (1991c), Int. J. Peptide Protein Res., **37**, 508-513

Qian, W., Bandekar, J., Krimm., S., (1991), Biopolymers, **31**, 193-210

Radford, S.E., Buck, M., Topping, K.D., Dobson, C.M., Evan, P.A., (1992), Proteins: Structure, Function and Genetics, **14**, 237-248

Rao, K.R., Brew, K., (1989), Biochem. Biophy. Res. Comm., **163**, 1390-1396

Rava, R.P., Spiro, T.G., (1985), Biochemistry, **24**, 1861-1865

Redfield, C.,Dobson, C.M., (1988), Biochemistry, **27**, 122-136

Roberts, G.M., Lee, O., Calienni, J., Diem, M., (1988), J. Am. Chem. Soc., **110**, 1749-1752

Rommel-Mohle, K., Hoffman, H.J., (1993), J. Mol. Structure (Theochem.), **285**, 211-219

Segawa, T., Sugai, S., (1983), J. Biochem., **93**, 1321-1328

Sekacis, I., Shenderovich, M., Nikiforovich, G., Liepins, E., Polevaya, L., Chipens, G., (1988), Coll. Czech Chem. Comm., **53**, 2810-2824

Siamwiza, M.N., Lord, R.C., Chen, M.C., Takamatsu, T., Harada, I., Maksuura, H., Shimanouch, T., (1975), Biochemistry, **14**, 4870- 4876

Simons, L., Bergstrom, G., Blomfelt, G., Forss, S., Stenback, H. Wansen, G., (1972), Soc. Scient. Fennica Comment. Physico-Math., **42**, 125-207

Smith, L.J., Sutcliffe, M.J., Redfield, R., Dobson, C.M., (1993a), J. Mol. Biol., **229**, 930-944

Smith, P.E., Pettitt, B.M., Karplus, M., (1993b), J. Phys. Chem., **97,** 6907-6913

Susi, H., Byler, D.M., Gerasimowicz, W.V., (1983), Journal Mol. Structure, **102**, 63-79

Tiffany, M.L., Krimm, S., (1968), Biopolymers, **6**, 1379-1382

Tobias, D.J., Brooks, C.L., (1992), J. Phys. Chem., **96**, 3864-3870

Tu, A., (1986), in Spectoscopy of Biological Systems, Advances in Spectroscopy, **13**, R.J.H. Clark and R.E.Hester (Eds), Wiley, Chichester, 47-112 (Chapter 2)

Urbanova, M., Dukor, R.K., Pancoska, P., Gupta, V.P., Keiderling, T.A., (1991), Biochemistry, **30**, 10479-10485

Velicelebi, G., Sturtevant, J.M., (1979), Biochemistry, **18**, 1180-1186

Wedin, R.E., Delepierre, M., Dobson, C.M., Poulsen, F.M., (1982), Biochemistry, **21**, 1098-1103

Wen, Z.Q., (1991), Doctoral Thesis, Glasgow University

Wen, Z.Q., Hecht, L., Barron, L.D., (1994a), J. Am. Chem. Soc., **116**, 443-445

Wen, Z.Q., Hecht, L., Barron, L.D., (1994b), Protein Science, **3**, 435-439

White, F.H., (1982), Biochemistry, **21**, 967-977

Williams, D.H., Fleming, I., (1989), Spectroscopic Methods in Organic Chemistry (4th ed.), McGraw-Hill, London

Wright, P.E., Dyson, H.J., Lerner, R.A., (1988), Biochemistry, **27**, 7167-7175

Yu, N.T., (1974), J. Am. Chem. Soc., **96**, 4664-4668

Yu, N.T., Jo, B.H., (1973), Arch. Biochem. Biophys., **156**, 469-474

Yutani, K., Ogasahara, K., Kuwajima, K., (1992), J. Mol. Biol., **228**, 347-350

Kettle, S.F.A., Lugwisha, E., Vorderwisch, P., Eckert, J., (1990), Spectrochim. Acta, **46A**, 921-926

Zimmermann, S.S., Pottle, M.S., Nemethy, G., Scheraga, H.A., (1977), Macromolecules, **10**, 1-9

Zuk, W.M., Freedman, T.B., Nafie, L.A., (1989), Biopolymers, **28**, 2025-2044

PUBLICATIONS

PAPERS AND COMMUNICATIONS.

Barron, L.D., Cooper, A., Ford, S.J., Hecht, L., Wen, Z.Q., (1992), Faraday Discuss., **93**, 259-268 (copy attached)

Ford, S.J., Wen, Z.Q., Hecht, L., Barron, L.D., (1994), Biopolymers, **34**, 303-313 (copy attached)

Barron, L.D., Ford, S.J., Bell, A.F., Wilson, G., Hecht, L., Cooper, A., (1994), Faraday Discuss., **99**, in press (copy attached)

Bell, A.F., Ford, S.J., Hecht, L., Wilson, G., Barron, L.D., (1994), Int. J. Biol. Macromol., **16**, 227-228 (copy attached)

Ford, S.J., Cooper, A., Hecht, L., Wilson, G., Barron, L.D., submitted to Faraday Transactions

Ford, S.J., Wilson, G., Cooper, A., Hecht, L., Barron, L.D., in preparation

CONFERENCE ABSTRACTS

Vibrational Raman optical activity of proteins. Ford, S.J. Wen, Z.Q., Hecht, L., Barron, L.D., In *Laser study of Macroscopic Biosystems* (ed. J.E.I. Korpii-Tommola), Proc. SPIE 1992, 1992, 46-50

Vibrational Raman optical activity of biological molecules. Barron, L.D., Hecht, L., Wen, Z.Q., Ford, S.J., In *Laser study of Macroscopic Biosystems* (ed. J.E.I. Korpii-Tommola), Proc. SPIE 1992, 1992, 2-11

Vibrational Raman optical activity of nucleosides and nucleotides. Hecht, L., Barron, L.D., Wen, Z.Q., Ford, S.J., Bell, A.F., In *Proceedings of the XIIIth International Conference on Raman Spectroscopy* (eds. W. Kiefer, M. Cardonna, G. Schaak, F.W. Schneider and H.W. Schrotter), Wiley, Chichester, 1992, 1098-1099

Barron, L.D., Hecht, L., Wen, Z.Q., Ford, S.J., In *Proceedings of the XIIIth International Conference on Raman Spectroscopy* (eds. W. Kiefer, M. Cardonna, G. Schaak, F.W. Schneider and H.W. Schrotter), Wiley, Chichester, 1992, 4642-4643

Vibrational Raman optical activity: from small organics to biopolymers. Barron, L.D., Hecht, L., Ford, S.J., Bell, A.F., Wen, Z.Q., In *Lectures and Posters of the 5th Inernational Conference on Circular Dichroism* (eds. K. Nakanishi, R.W. Woody and N. Berova), Colorada State University, Fort Collins, 1993, 158-63