SIMULATIONS AND ANALYSIS OF THE RAMAN SCATTERING AND DIFFERENTIAL RAMAN SCATTERING/RAMAN OPTICAL ACTIVITY (ROA) SPECTRA OF AMINO ACIDS, PEPTIDES AND PROTEINS IN AQUEOUS SOLUTION

K. J. Jalkanen*, R. M. Nieminen*, and J. Bohr**

The Raman and Raman optical activity (ROA) spectra of amino acids and small peptides in aqueous solution have been simulated by density functional theory and restricted Hartree–Fock methods. The treatment of the aqueous environment in treated in two ways. The water molecules in the first hydration shell which strongly interact with the molecule are treated explicitly while the waters in the bulk are treated by a continuum model. The structures are optimized and the harmonic force fields are calculated. The derivatives needed to simulate the Raman and ROA intensities are calculated from first principles. The simulated Raman and ROA spectra have been compared to recently meassured spectra on amino acids and peptides. The simulations and understanding from them are used to interpret the Raman and ROA spectra of proteins. A comparison to vibrational absorption (VA) and vibrational circular dichroism (VCD) spectra is also given, and the complementarity of the information gained from the Raman and ROA spectra with that gained from VA and VCD spectra is discussed.

Introduction

Raman optical activity measurements (ROA) [1] and calculations [2] have recently been applied to systems of biological interest. Raman and ROA spectroscopy have the advantage over vibrational absorption (VA) and vibrational circular dichroism (VCD) for biological samples because water is not a strong Raman scatterer but a strong infrared absorper. The alkalii halide windows normally used in IR and VCD studies on molecules in nonpolar solvents are not inert to aqueous solutions used for biosamples. Use of CaF_2 windows leads to a loss of information due to the IR cutoff of this material. For Raman scattering quartz windows can be used.

Our interest in Raman and ROA spectra arises from the need to know the secondary structure of proteins, peptides and other biomolecules. Currrently the only routine methods to determine the structure of a protein are X-Ray crystallography and NMR spectroscopy. For an X-ray structure determination, one requires crystals and hence to be able to crystallise the protein. In NMR structure measurements one needs enough assigned nuclear Overhauser peaks (NOE's) to uniquely determine the structure. Any flucuations within the protein on the NMR scale will be averaged. Also both the X-ray crystallography and NMR methods only determine the global minimum and one has problems when one has multiple conformers present. Which structure is the one that is biologically active is a question which is open to debate. Raman and ROA spectrocopy are in principal able to be applied to higher energy states and also to multiple conformers, assuming that one is able to resolve and assign the bands to the various conformers. What has held back the application of Raman and ROA spectroscopy to structure determination has been an adequate theory. In a recent review on ROA Barron stated that hirthereto *ab initio* calculations had not been developed sufficiently to assist in the interpretation of the observed ROA spectra of the biomolecules [3]. In his review recently measured Raman and ROA spectra of a series of proteins, disaccharides and nucleic acids are presented. Recently we have simulated the VA and VCD spectra of NALANMA and used the predictions to interpretate the VCD spectra of proteins [4]. Here we will also interpret the ROA spectra of proteins based on our ROA simulations of NALANMA [5]. This molecule, capped L-alanine, has been used to test various empirical force fields, semi-empirical methods and *ab initio* methds, for adequacy and accuracy in modelling biomolecules. The calculations of Raman and ROA intensities requires a structure determination, at this local minimum structure a Hessian to determine the vibrational frequencies and atomic displacements, and the polarizability derivatives. In this sense, Raman and ROA are very closely related to VA and VCD spectroscopy. The difference between VA and VCD and Raman and ROA are in the theory needed to predict the intensities at the given frequencies. Upon molecular conformational changes the vibrational frequencies as well as the intensities change. If one were just to use the frequencies any of VA, VCD, Raman or ROA would give this information, under the assumption that one was able to assign the modes and resolve the individual spectral lines or at least deconvolute them. If the information in the vibrational frequencies is not sufficient to determine conformation, then one can calculate the VA, VCD, Raman and/or ROA spectral intensities or a combination of them and use this combined information to give us the information we need to determine conformation.

* Laboratory of Physics, Helsinki University of Technology, P.O. Box 1100, Otakaari 1 M, FIN-02015 HUT, Finland.
** Department of Physics, The Technical University of Denmark, Building 307, DK-2800 Lynbgy, Denmark.

Methods

Raman calculations

The calculation of the Raman scattering intensities requires the calculation of the electric dipole-electric dipole polarisability derivatives (EDEDPD):

$$\alpha_{ij}^{\lambda\alpha} = \left(\frac{\partial^3}{\partial X_{\lambda\alpha}\partial E_i \partial E_j} W_G\left(\overline{R}, E_i, E_j\right)\right)_{\overline{R}_e, E_j=0, E_j=0}$$

$$= \left(\frac{\partial}{\partial X_{\lambda\alpha}} \alpha_{i,j}\left(\overline{R}\right)\right)_{R_e} \tag{1}$$

Amos [6] and Frisch *et al* [7]. has implemented the analytical calculation of the EDEDPD at the RHF level. Recently Johnson and Florian [8] have implemented the polarisability derivatives at the DFT level of theory using the finite-field numerical second differentiation of analytically calculated energy derivatives. Subsequently Gaussian 98 has also implemented the calculation of the EDEDPD at the DFT level. This is the same method used by Komornicki and McIver to calculate Raman intensities by Hartree-Fock theory in their program GRADSCF [9]. After the EDEDPD have been calculated they can be combined with the vibrational frequencies and normal mode vectors to predict the differential Raman scattering cross sections, scattering activities and depolarisation ratios. The absolute differential Raman scattering cross sections is given by

$$\left(\frac{d\sigma_i}{d\Omega}\right) = \frac{2^4 \prod^4}{45} \frac{\left(\nu_0 - \nu_j\right)^4}{1 - \exp\left[-\frac{hc\nu_j}{kT}\right]} \frac{h}{8\prod^2 c\nu_j} S_j, \tag{2}$$

where $\nu\nu_0$ is the exciting frequency, h, c and k are Plank's constant, the speed of light and Boltzmann's constant, S_j is the Raman scattering activity, given by

$$S_j = g_j\left(45\overline{\alpha}_j^2 + 7\overline{\beta}_j^2\right), \tag{3}$$

where g_j is the degeneracy of the ith transition of energy $h\nu\nu_i$ and $\overline{\alpha}_j^2$ and $\overline{\beta}_j^2$ are given by

$$\overline{\alpha}_j^2 = \frac{1}{9}\left(\alpha_{xx}^j + \alpha_{yy}^j + \alpha_{zz}^j\right)^2 \tag{4}$$

$$\overline{\beta}_j^2 = \frac{1}{2}\Big\{\left(\alpha_{xx}^j - \alpha_{yy}^j\right)^2 + \left(\alpha_{xx}^j - \alpha_{zz}^j\right)^2 + \left(\alpha_{yy}^j - \alpha_{zz}^j\right)^2 + 6\left[\left(\alpha_{xy}^j\right)^2 + \left(\alpha_{yz}^j\right)^2 + \left(\alpha_{xz}^j\right)^2\right]\Big\}$$

Raman optical activity calculations

The calculation of the Raman optical activity (ROA) scattering intensities requires the calculation of the electric dipole-electric dipole, electric dipole-electric quadrupole and electric dipole-magnetic dipole polarisability derivatives. These tensors can be related to the electric dipole, μ, magnetic dipole, m, and electric quadrupole, θ, moments by the following equations:

$$\mu_\alpha = \mu_\alpha^o + \alpha_{\alpha\beta}F_\beta + \frac{1}{3}A_{\alpha\beta\gamma}F_{\beta\gamma} + \frac{1}{\omega}G_{\alpha\beta}\dot{B}_\beta + \ldots \tag{5}$$

$$m_\alpha = m_\alpha^o - \frac{1}{\omega}G_{\beta\alpha}\dot{F}_\beta + \ldots \tag{6}$$

$$\theta_{\alpha\beta} = \theta_{\alpha\beta}^o + A_{\gamma\alpha\beta}F_\gamma + \ldots, \tag{7}$$

where the F_α and B_β are the static electric and magnetic fields respectively and $\dot{F}_\beta$ and $\dot{B}_\beta$ are the dynamic electric and magnetic fields respectively and $F'_{\alpha\beta}$ is the electric field gradient [10].

From the above equations we can take various energy derivatives to get the electric dipole-electric quadrupole polarisability derivative (EDEQPD) as a third derivative of the energy or simple as a first derivative,

$$A_{ijk}^{\lambda\alpha} = \left(\frac{\partial^3}{\partial X_{\lambda\alpha}\partial F_i \partial F'_{jk}} W_G\left(\overline{R}, F_i, F'_{jk}\right)\right)_{\overline{R}_e, F_i=0, F'_{jk}=0}$$

$$= \left(\frac{\partial}{\partial X_{\lambda\alpha}} A_{ijk}\left(\overline{R}\right)\right)_{R_e}. \tag{8}$$

Similarly one can determine the electric dipole-magnetic dipole polarisability derivative (EDMDPD) as a third derivative of the energy or as a simple first derivative

$$G_{ij}'^{\lambda\alpha} = \left(\frac{\partial^3}{\partial X_{\lambda\alpha}\partial F_i \partial \dot{B}_j} W_G\left(\overline{R}, F_i, \dot{B}_j\right)\right)_{\overline{R}_e, F_i=0, \dot{B}_j=0}$$

$$= \left(\frac{\partial}{\partial X_{\lambda\alpha}} G'_{ij}\left(\overline{R}\right)\right)_{R_e}. \tag{9}$$

Multiplying the Cartesian polarisability derivatives by the $S_{\lambda\alpha,i}$ we can get the polarisability derivatives with respect to normal coordinates. The following quantities can then be calculated which will be useful for expressing the various ROA intensities which one get for the various experimental set-ups of the ROA experiment,

$$\overline{\alpha}_j\overline{G}'_j = \frac{1}{9}\left(\alpha_{xx}^j + \alpha_{yy}^j + \alpha_{zz}^j\right)\left(G'^{\,j}_{xx} + G'^{\,j}_{yy} + G'^{\,j}_{zz}\right), \tag{10}$$

$$\begin{aligned}\gamma_j^2 = \Big\{&\left(\alpha_{xx}^j - \alpha_{yy}^j\right)\left(G'^{\,j}_{xx} - G'^{\,j}_{yy}\right) + \left(\alpha_{xx}^j - \alpha_{zz}^j\right)\\ &+ \left(G'^{\,j}_{xx} - G'^{\,j}_{zz}\right) + \left(\alpha_{yy}^j - \alpha_{zz}^j\right)\left(G'^{\,j}_{yy} - G'^{\,j}_{zz}\right)\\ &+3\Big[\alpha_{xy}^j\left(G'^{\,j}_{xy} + G'^{\,j}_{yx}\right) + \alpha_{xz}^j\left(G'^{\,j}_{xz} + G'^{\,j}_{zx}\right)\\ &+\alpha_{yz}^j\left(G'^{\,j}_{yz} + G'^{\,j}_{zy}\right)\Big]\Big\},\end{aligned} \tag{11}$$

$$\begin{aligned}\delta_j^2 = \frac{\omega}{2}\Big\{&\left(\alpha_{yy}^j - \alpha_{xx}^j\right)A_{zxy}^j + \left(\alpha_{xx}^j - \alpha_{zz}^j\right)A_{yzx}^j\\ &+\left(\alpha_{zz}^j - \alpha_{yy}^j\right)A_{xyz}^j\\ &+\alpha_{xy}^j\left(A_{yyz}^j - A_{zyy}^j + A_{zxx}^j - A_{xxz}^j\right)\\ &+\alpha_{xz}^j\left(A_{yzz}^j - A_{zzy}^j + A_{xxy}^j - A_{yxx}^j\right)\\ &+\alpha_{yz}^j\left(A_{zzx}^j - A_{xzz}^j + A_{xyy}^j - A_{yyx}^j\right)\Big\}.\end{aligned} \tag{12}$$

We have calculated numerically the electric dipole-electric dipole polarisability derivatives, the electric dipole-electric quadrupole polarisability derivatives and the electric dipole-magnetic dipole polarisability derivatives. We

Measurement type	$I_R - I_L$	$I_R + I_L$
Depolarised	$\frac{12\omega}{c}\left(\frac{\gamma_j^2}{\omega} - \frac{1}{3}\frac{\delta_j^2}{\omega}\right)$	$6\beta_j^2$
Polarised	$\frac{2\omega}{c}\left(\frac{45}{\omega}\overline{\alpha}_j\overline{G}'_j + \frac{7\gamma_j^2}{\omega} + \frac{\delta_j^2}{\omega}\right)$	$45\overline{\alpha}_j^2 + 7\beta_j^2$
Magic angle	$\frac{20\omega}{3c}\left(\frac{9}{\omega}\overline{\alpha}_j\overline{G}'_j + \frac{2\gamma_j^2}{\omega}\right)$	$\frac{10}{3}\left(9\overline{\alpha}_j^2 + 2\beta_j^2\right)$
Backward	$\frac{48\omega}{c}\left(\frac{\gamma_j^2}{\omega} + \frac{1}{3}\frac{\delta_j^2}{\omega}\right)$	$2\left(45\overline{\alpha}_j^2 + 7\beta_j^2\right)$
Forward	$\frac{8\omega}{c}\left(\frac{45}{\omega}\overline{\alpha}_j\overline{G}'_j + \frac{\gamma_j^2}{\omega} - \frac{\delta_j^2}{\omega}\right)$	$2\left(45\overline{\alpha}_j^2 + 7\beta_j^2\right)$

have used two point finite differencing and with a displacement of 0.005 Å [5].

These quantities can be combined to give the following intensities for the five Raman optical activity experimental measurement types [11], where I_R and I_L are the Raman scattered intensities with linear α-polarisation in right and left circularly polarised incident light (see the table).

Results and Discussion

Recently Barron *et al.* 2000 has reported the Raman and ROA spectra for proteins for which the X-ray structures have been determined. They have characterized the α-helical structure by its amide I ROA couplet centered at approximately 1650 cm^{-1}, positive ROA intensity in the range from 870 to 950 cm^{-1} and tentative assignments of ROA intensities in the 1290 to 1340 cm^{-1} region based on the degree of hydration of the helix or whether the helix is in a hydrophobic environment. The β-sheet structure has been characterized by negative ROA band in the region 1230 to 1250 cm^{-1}, an amide I ROA couplet centered at approximately 1655 to 1670 cm^{-1}. It would be nice to see what the predictions can be made for the α-helical and β-sheet structure. The simplest model system which has been used to model the protein backbone structure has been capped L-alanine, also called N-acetyl L-alanyl N′-methylamide (NALANMA). We have previously modeled various structures of solvated NALANMA to determine its structure in aqueous medium [4] and compared it to the structure of NALANMA in the isolated (non hydrated) state [5]. Here we compare the model calculations for NALANMA with the measured protein spectra. One thing to note is that the ROA spectra of protein structures do not seem to be as rich as those of simple peptides. This is probably due to to unresolved features. The significance of individual versus collective features awaits to be investigated. A calculation of the resulting spectra is necessary for its interpretation. Here the calculation of the ROA spectra for the variety of different local conformers present within the protein must be undertaken. Then the effect of environment must also be taken into account. Barron and coworkers have pointed out that the degree of hydration and also the hydrophobic environment for the α-helix structures appears to give different ROA spectra. This underlies the importance of modelling the environment. Previous Raman and ROA and also VA and VCD simulations have found good agreement between calculated and measured spectra. But a majority of the comparisons have been made for spectra measured for the molecules in nonpolar solvents. This was an important first step to document the accuracy of the theory. The molecules also had little if any conformational flexibility, that is, most had one local miniumum which was much lower in energy than the other possible structures. But recent work on NALANMA and LA has shown that the VA, VCD, Raman and ROA spectra change when one changes the solvent from either carbon tetrachloride or chloroform to water. These spectral (and hence conformational/structural) changes have been simulated (albeit with much more work).

To simulate the VA and VCD spectra one needs to calculate the atomic polar tensors (APT) and atomic axial tensors (AAT) [12]. Here we measure electric dipole mediated transitions and magnetic dipole mediated transitions. Hence we gain different information from the VA and VCD intensities than we do from the Raman and ROA intensities. Recently experimental work has been done to measure the VA and VCD spectra of amino acids, peptides and proteins in aqueous solution [13]. The complementary nature and experimental setup comparing ROA and VCD has also recently been reviewed [14]. The VCD spectra has been measured mostly in the amide I region (the C=O stretch), the same region where people have used VA spectra to determine the percentage of various secondary structual elements present in proteins [15]. Similarly one can use Raman scattering spectra to get information about secondary structural elements in proteins [16]. What makes VCD and ROA attractive is that like CD in the electronic spectra, the bisignate nature of the VCD and ROA spectra allow for in many cases better resolution of the spectra and hence the interpretation is less dependent on the assumed number of bands under a very broad absorption peak. We have also shown recently that the amide I intensities are very dependent on the conformer, the α-helical amide I intensities are much larger than they are for other secondary structural elements. Further it has been show that proteins which have similar CD spectra can have different VCD spectra. Much work on the problem of assigning the differences to specific secondary structural elements remains. Our simulations of the VCD and ROA spectra of NALANMA, L-alanine, and L-alanyl-L-alanine in aqueous

solution are a first step in providing a sound and rigorous interpretation of the experimentally measured VCD and ROA spectra. Hence our work has confirmed the bisignate nature of the amide I ROA bands for the α-helical structure and β-sheet type structure. In the isolated state of NALANMA there are two distinct conformers, the C_7^{eq} and structures, which collapse to a single minimum, the P_{II} structure. The agreement between the calculated Raman and ROA spectra for NALANMA in aqueous is in very good agreement with experimentally reported specta [5]. We look forward to presenting a more complete interpretation of the recently presented ROA spectra of proteins of Barron in a future publication based on *ab initio* simulations. Previous interpretations of the ROA spectra of proteins were based solely on trying to correlate the secondary structural elements of proteins for which the X-ray structures have been determined with their measured ROA spectra. The time now has come to supplement this work with the simulated Raman and ROA spectra of amino acids and peptides in the hydrated and non-hydrated states. Similar work has been performed for VA and VCD. The method of using structural correlations derived from VA and VCD spectra of proteins with X-ray determined with theoretical simulations seems to be the best way to interpret the VA, VCD, Raman and ROA spectra of proteins and other large biomolecules. The use of either method on its own is insufficient, as shown by the reassignment of modes based on using only correlations between spectra of proteins with known structures. The use of theoretical simulations adds much to the interpretation of the spectra, and one is also able to calculate the VA, VCD, Raman and ROA spectra for nonnative states and then to interpret the spectra of proteins in their nonnative state(s).

Acknowledgements

KJJ would like to thank the Finnish Academy of Science for his scholarship for his stay at HUT.

References

1. L. Hecht and L.D. Barron, *Faraday Discuss.* **99** (1994) 35; L.D. Barron, S.J. Ford, A.F. Bell, G. Wilson, L. Hecht, and A. Cooper, *Faraday Discuss.* **99** (1994) 217.
2. T. Helgaker, K. Ruud, K.L. Bak, P. Jørgensen, and J. Olsen, *Faraday Discuss.* **99** (1994) 165; P.L. Polavarapu and Z. Deng, *Faraday Discuss.* **99** (1994) 151.
3. L.D. Barron, L. Hecht, E.W. Blanch, and A.F. Bell, *Progress in Biophysics and Molecular Biology*, In Press.
4. K.J. Jalkanen and S. Suhai, *Chem. Phys.* **208** (1996) 81; Z. Deng, P.L. Polavarapu, S.J. Ford, L.D. Barron, C.S. Ewig, and K. Jalkanen, *J. Phys. Chem.* **100** (1996) 2025.
5. W.-G. Han, K.J. Jalkanen, M. Elstner, and S. Suhai, *J. Phys. Chem.* **B 102** (1998) 2587.
6. R.D. Amos, *Chem. Phys. L.* **124** (1986) 376; R.D. Amos in *Ab Initio Methods in Quantum Chemistry*; Lawley, K.P.; Ed.; Wiley: New York, 1987.
7. M.J. Frisch, Y. Yamaguchi, J.F. Gaw, H.F. Schaefer III, and J.S. Binkley, *J. Chem. Phys.* **84** (1986) 531.
8. B.G. Johnson and J. Florian, *Chem. Phys.* **L. 247** (1995) 120–125.
9. A. Kormonicki and J.W. McIver Jr., *J. Chem. Phys.* **70** (1979) 2014.
10. A.D. Buckingham, *Adv. Chem. Phys.* **12** (1967) 107.
11. L.D. Barron, *Molecular Light Scattering and Optical Activity.* Cambridge University Press: Cambridge, UK 1982.
12. R.D. Amos, *Chem. Phys. L.* 108 (1984) 185; P.J. Stephens, *J. Phys. Chem.* **89** (1985) 748; A.D. Buckingham, P.W. Fowler, and P.A. Galwas, *Chem. Phys.* **112** (1987) 1.
13. P. Xie, Q. Zhou, and M. Diem, *Faraday Discuss.* **99** (1994) 233; P. Pancoska, E. Bitto, V. Janota, and T.A. Keiderling, *Faraday Discuss.* **99** (1994) 287.
14. L.A. Nafie, G.-S. Yu, X. Qu, and T.B. Freedman, *Faraday Discuss.* **99** (1994) 13; L.A. Nafie, *Applied Spectrosc.* **50** (1996) 14A.
15. J. Contzen and C. Jung, *Biochemistry* **37** (1998) 4317; C. Mouro, C. Jung, A. Bondon, and G. Simonneaux, *Biochemistry* **36** (1997) 8125; N.N. Kalnin, I.A. Baikalov, and S.Yu. Venyaminov, *Biopolymers* **30** (1990) 1273; S.Yu. Venyaminov and N.N. Kalnin, *Biopolymers* **30** (1990) 1243; P.I. Haris and D. Chapman, *Biopolymers (Peptide Science)* **37** (1995) 251.
16. L.G. Tensmeyer and E.W. Kauffman, *II in Spectroscopic Methods for Determining Protein Structure in Solution*, 1996, Chapter 5, pp. 69–95, VCH Verlagsgesellshaft mbH, Weinheim, Germany.